



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 114883**

**TO: Jeffrey Parkin**  
**Location: rem/3e39/3c18**  
**Art Unit: 1648**  
**February 27, 2004**

**Case Serial Number: 09/891609**

**From: P. Sheppard**  
**Location: Remsen Building**  
**Phone: (571) 272-2529**

**sheppard@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**

STIC-Biotech/ChemLib

114883

From: Parkin, Jeffrey  
Sent: Sunday, February 22, 2004 4:51 PM  
To: STIC-Biotech/ChemLib  
Subject: U.S. Serial No. 09/891,609

Please search **SEQ ID NOS.: 2 and 4** from the aforementioned application (09/891,609; Stamatatos et al.) v. all relevant databases, including interference. Please provide a paper copy and electronic copy of the results.

Thanks!

JSP  
AU 1648  
REM 3E39  
2-0908

RECEIVED  
FEB 23 2004  
STIC

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**This Page Blank (uspto)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:02:44 ; Search time 55.1526 Seconds  
(without alignments)  
3309.468 Million cell updates/sec

Title: US-09-891-609A-2  
Perfect score: 3474  
Sequence: 1 MRVGIKKNYQHLWRGTTLL.....DKWASLWNWFDISKWLWYIK 646

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3474	100.0	646	5	Aau75155 Modified
2	3422	98.5	847	3	Aay97073 Variant H
3	3327	95.8	619	5	Aau75156 N-termina
4	3279.5	94.4	842	6	Abr55684 HIV isola
5	3279.5	94.4	842	7	Adcl13218 Protein o
6	3274.5	94.3	842	6	Abu66565 Human Imm
7	3264.5	94.0	842	5	Abb06211 HIV Env 1
8	3065	88.2	883	4	Aab82761 Ancestral
9	3062.5	88.2	850	2	Aar67724 gp120 fro
10	3006.5	86.5	856	6	Abr55495 Amino aci
11	2997	86.3	851	1	Aap80967 HIV prote
12	2984	85.9	855	2	Aaw53112 ENV prote
13	2984	85.9	855	3	Aay77298 HIV-1 (AT
14	2984	85.9	855	3	Aay77302 HIV-1 (AT
15	2984	85.9	855	6	Abu57553 AIDS asso
16	2984	85.9	855	6	Abu57550 AIDS asso
17	2984	85.9	855	6	Abu63182 Protein #
18	2984	85.9	855	6	Abu63186 Env prote
19	2984	85.9	863	1	Aap61509 Sequence
20	2984	85.9	863	2	Aar29706 env gene
21	2984	85.9	863	5	Aae35790 HIV-2 (9B
22	2978.5	85.7	855	2	Aawl1581 Human Imm
23	2978.5	85.7	855	2	Aaw88113 Env prote
24	2974	85.6	857	2	Aar67725 gp120 fro
25	2951.5	85.0	860	2	Aaw31284 HIV-SF2 v

26	2940.5	84.6	855	2	AAW43069	Aaw43069 HIV-1 gpl
27	2939.5	84.6	856	2	AAR67726	Aar67726 gp120 fro
28	2939.5	84.6	856	4	AAB85999	Aab85999 Amino aci
29	2939.5	84.6	863	2	AAR43869	Aar43869 HTLV-III
30	2933	84.4	868	5	AAO19389	Aao19389 Lymphaden
31	2931.5	84.4	856	2	AAR25940	Aar25940 Modified
32	2931.5	84.4	866	1	AAP80966	Aap80966 HIV prote
33	2930	84.3	868	1	AAP60422	Aap60422 Sequence
34	2929.5	84.3	863	1	AAP60349	Aap60349 HTLV-III
35	2926.5	84.2	856	1	AAP61514	Aap61514 Sequence
36	2926.5	84.2	856	2	AAW89325	Aaw89325 HIV-1 env
37	2926.5	84.2	856	6	ABU63322	Abu63322 Human lym
38	2926	84.2	868	1	AAP60063	Aap60063 HIV virus
39	2925	84.2	865	1	AAAP70175	Aap70175 Sequence
40	2922.5	84.1	856	3	AAAY97072	Aay97072 Wild type
41	2922.5	84.1	856	8	ADE84721	Ade84721 Human imm
42	2919.5	84.0	854	3	AAAB10053	Aab10053 HIV-1 coa
43	2919.5	84.0	854	3	AAAB10697	Aab10697 HIV-1 env
44	2919.5	84.0	854	4	AAAB86199	Aab86199 HIV gp41
45	2919.5	84.0	854	5	ABG73663	Abg73663 HIV-1 NL4

ALIGNMENTS

RESULT 1  
AAU75155  
ID AAU75155 standard; protein; 646 AA.  
XX  
AC AAU75155;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Modified full-length HIV-1 SF162DV2 gp140 envelope protein.  
XX  
KW Human immunodeficiency virus type 1; HIV-1; gp140 envelope protein;  
KW second hypervariable region; V2; SF162deltaV2; SF162DV2; AIDS;  
KW acquired immunodeficiency syndrome; anti-HIV; mutant; mutein; immunogen.  
XX  
OS Human immunodeficiency virus 1; strain SF162 (clade B).  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 539  
FT /note= "Encoded by CAGCTC"  
XX  
PN WO200200250-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 27-JUN-2001; 2001WO-US020483.  
XX  
PR 27-JUN-2000; 2000US-0214608P.  
PR 26-JUN-2001; 2001US-00891609.  
XX  
PA (AARO-) AARON DIAMOND AIDS RES CENT.  
XX  
PI Stamatos L, Barnett S, Shrivastava I;  
XX  
DR WPI; 2002-130836/17.  
XX  
DR N-PSDB; AAS15498.  
XX  
PT Immunizing an animal and eliciting an immune response against  
PT heterologous HIV-1 in an animal, involves administering an immunogen  
PT comprising modified HIV-1 envelope protein, or DNA or virus encoding the  
PT protein.  
XX  
XX Claim 5; Fig 16; 62pp; English.  
XX  
XX The present invention relates to methods for immunising an animal,  
XX preferably human, against heterologous human immunodeficiency virus type  
XX 1 (HIV-1), and eliciting a heterologous immune response to HIV-1 in that  
XX animal. The method comprises administering an immunogen having at least

CC one modified HIV-1 envelope protein with a deletion in the second  
CC hypervariable (V2) region or its fragment, or DNA or virus encoding the  
CC modified HIV-1 envelope protein. The invention discloses the  
CC polynucleotide and amino acid sequences for the full-length HIV-1  
CC SF162deltaV2 (SF162DV2) gp140 envelope protein, and, for an N-terminal  
CC mutant of SF162DV2. The immunogens of the invention are useful for  
CC immunising an animal against heterologous HIV-1 strains by eliciting  
CC neutralising antibodies or protective antibodies in the animal. The  
CC method can be used to treat acquired immunodeficiency syndrome (AIDS).  
CC The present sequence represents the full-length modified HIV-1 SF162DV2  
CC gp140 envelope protein  
XX  
SQ Sequence 646 AA;  
  
Query Match 100.0%; Score 3474; DB 5; Length 646;  
Best Local Similarity 100.0%; Pred. No. 1.9e-169; Indels 0; Gaps 0;  
Matches 646; Conservative 0; Mismatches 0;  
  
QY 1 MRVKGIRKXQHLWRGGTLLGLMLCSAVEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
DB 1 MRVKGIRKXQHLWRGGTLLGLMLCSAVEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
  
QY 61 DTEVHNWATHACVPTDNPQBIIVLENTNFNMKNMVEQMHEDIISLWQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDNPQBIIVLENTNFNMKNMVEQMHEDIISLWQSLKPCVK 120  
  
QY 121 LTPCLVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSFKVGAGKLINCNTSVITQCPKVS 180  
DB 121 LTPCLVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSFKVGAGKLINCNTSVITQCPKVS 180  
  
QY 181 FEPITHYCAPAGFAILKCNKDKFNGSGCTNVTSTVQCTHGRPVVSTOLLNGSLAERG 240  
DB 181 FEPITHYCAPAGFAILKCNKDKFNGSGCTNVTSTVQCTHGRPVVSTOLLNGSLAERG 240  
  
QY 241 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPCRAFYATGDIIGDIRQAH 300  
DB 241 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPCRAFYATGDIIGDIRQAH 300  
  
QY 301 CNISGEKWNNTLKQIVTKLQAOFGNKTIIVFKOSSGGDPEIVMHSFNCGGEPFFYCNSTOLF 360  
DB 301 CNISGEKWNNTLKQIVTKLQAOFGNKTIIVFKOSSGGDPEIVMHSFNCGGEPFFYCNSTOLF 360  
  
QY 361 NSTWNNTIGPNTNGTITLPCRIKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRD 420  
DB 361 NSTWNNTIGPNTNGTITLPCRIKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRD 420  
  
QY 421 GGKEISNTTEIFRPGGGDRDNWRSELYKYKVVKIEPLGVAPTAKRRVYVQREKRAVTLG 480  
DB 421 GGKEISNTTEIFRPGGGDRDNWRSELYKYKVVKIEPLGVAPTAKRRVYVQREKRAVTLG 480  
  
QY 481 AMFLGFLGAAGSTMGARSITLTVOARQLISGIVQQQNLLRAIEAQOHLQLTVWGIKLO 540  
DB 481 AMFLGFLGAAGSTMGARSITLTVOARQLISGIVQQQNLLRAIEAQOHLQLTVWGIKLO 540  
  
QY 541 ARVLAVERYLKDQQLLIGWCSGLICTTAVPNASWSNKSLLDQIWNNTMWEWEIREIDN 600  
DB 541 ARVLAVERYLKDQQLLIGWCSGLICTTAVPNASWSNKSLLDQIWNNTMWEWEIREIDN 600  
  
QY 601 YTNLIYTLIEESQNOQEKNEQELLELDKWASLWNNWFDISKWLWYIK 646  
DB 601 YTNLIYTLIEESQNOQEKNEQELLELDKWASLWNNWFDISKWLWYIK 646  
  
RESULT 2  
ID AAY97073  
AC AAY97073 standard; protein; 847 AA.  
XX AAY97073;  
XX  
DT 12-SEP-2003 (revised)  
DT 31-OCT-2000 (first entry)  
XX  
XX Variant HIV-1 SF162 Env gp160.

XX HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;  
KW CD4 binding region; V1/V2 loop; bridging sheet.  
XX Human immunodeficiency virus 1; isolate SF162.  
XX WO200039303-A2.  
XX 06-JUL-2000.  
XX 30-DEC-1999; 99WO-US031272.  
XX 31-DEC-1998; 98US-0114495P.  
XX 29-SEP-1999; 99US-0156670P.  
XX (CHIR ) CHIRON CORP.  
XX Barnett S, Hartog K, Martin E;  
XX WPI; 2000-465745/40.  
XX Novel modified HIV Env polypeptides useful as immunizing agents and for  
XX preparing a vaccine to elicit an immune response against a broad range of  
XX HIV subtypes.  
XX Claim 5; Page 115-117; 139pp; English.  
XX Novel immunogenic modified human immunodeficiency virus (HIV) envelope  
XX (Env) polypeptides having an amino acid deleted or replaced in the region  
XX corresponding to residues 420-436 or 119-123 and 199-210 relative to  
XX isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV  
XX strain SF162, with numbering relative to isolate HXB-2. The Env  
XX polypeptides are modified so as to expose at least part of the CD4  
XX binding region. The modified HIV Env polypeptides, coding polynucleotides  
XX and constructs, further comprising an adjuvant, are used for inducing an  
XX immune response in an individual. The method involves administering a  
XX first composition comprising a polynucleotide encoding the Env  
XX polypeptide in a priming step and administering a second composition  
XX comprising a modified Env polypeptide as a booster in an amount  
XX sufficient to induce an immune response in the individual. The first  
XX and/or second composition further comprises an adjuvant (claimed). The  
XX intracellularly produced Env polypeptides can be used for a number of  
XX diagnostic and therapeutic purposes to determine the presence of reactive  
XX antibodies and/or Env proteins in a biological sample to aid in the  
XX diagnosis of HIV infection or disease status or as measure of response to  
XX immunization. (Updated on 12-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 847 AA;  
  
Query Match 98.5%; Score 3422; DB 3; Length 847;  
Best Local Similarity 95.4%; Pred. No. 1.1e-166; Indels 28; Gaps 2;  
Matches 643; Conservative 0; Mismatches 3;  
  
QY 1 MRVKGIRKXQHLWRGGTLLGLMLCSAVEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
DB 1 MRVKGIRKXQHLWRGGTLLGLMLCSAVEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
  
QY 61 DTEVHNWATHACVPTDNPQBIIVLENTNFNMKNMVEQMHEDIISLWQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDNPQBIIVLENTNFNMKNMVEQMHEDIISLWQSLKPCVK 120  
  
QY 121 LTPCLVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSFKVGAG----- 162  
DB 121 LTPCLVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSFKVTTTIRNKKQKEYALFYKLDVV 180  
  
QY 163 -----KLINCNTSVITQCPKVSFEPPIHYCAPAGFAILKCNKDKFNGSGPCTNV 213  
DB 181 PIDNDNTSYKLINCNTSVITQCPKVSFEPPIHYCAPAGFAILKCNKDKFNGSGPCTNV 240  
  
QY 214 STVQCTHGIRPVVSTOLLNGSLAERGVIIRSEFTDNAKTIIVQLKESVEINCTRPNNN 273  
DB 241 STVQCTHGIRPVVSTOLLNGSLAERGVIIRSEFTDNAKTIIVQLKESVEINCTRPNNN 300

QY 274 TRKSTIGRGRAYATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLQAFGNKTIIVFKQS 333  
Db 301 TRKSTIGRGRAYATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLQAFGNKTIIVFKQS 360  
QY 334 SGSDPEIVMHSFNCGGEFFYCNSTQLFNSTWNTTIGPNNNTGTITLPCRIKQIINRWQEV 393  
Db 361 SGSDPEIVMHSFNCGGEFFYCNSTQLFNSTWNTTIGPNNNTGTITLPCRIKQIINRWQEV 420  
QY 394 GKAMVAPPVIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSELKYKVV 453  
Db 421 GKAMVAPPVIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSELKYKVV 480  
QY 454 KIEPLGVAPTKARRVQREKRAVTLGAMFLGAGSTMGARSITLTVQARQLLSGIV 513  
Db 481 KIEPLGVAPTKARRVQREKRAVTLGAMFLGAGSTMGARSITLTVQARQLLSGIV 540  
QY 514 QOONLLRAIEAQHLLQITVMGK-LQARVLAVERYLKDQQLGWTGCGSKLICTTAVP 572  
Db 541 QOONLLRAIEAQHLLQITVMGK-LQARVLAVERYLKDQQLGWTGCGSKLICTTAVP 600  
QY 573 WNASNSKSLDQIWNMTWMEWEREIDNYTNLIYTLIEESQOQEKNEQELLELOKWASL 632  
Db 601 WNASNSKSLDQIWNMTWMEWEREIDNYTNLIYTLIEESQOQEKNEQELLELOKWASL 660  
QY 633 WNWFDISKWLWYIK 646  
Db 661 WNWFDISKWLWYIK 674

RESULT 3

AAU75156  
ID AAU75156 standard; protein; 619 AA.  
XX  
AC AAU75156;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE N-terminal mutant of modified HIV-1 SF162DV2 gp140 envelope protein.  
XX  
KW Human immunodeficiency virus type 1; HIV-1; gp140 envelope protein;  
KW second hypervariable region; V2; SF162deltaV2; SF162DV2; AIDS;  
KW acquired immunodeficiency syndrome; anti-HIV; mutant; mutein; immunogen.  
XX  
OS Human immunodeficiency virus 1; strain SF162 (clade B).  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 512 /note= "Encoded by CAGCTC"  
FT  
XX  
PN WO200200250-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 27-JUN-2001; 2001WO-US020483.  
XX  
PR 27-JUN-2000; 2000US-0214608P.  
PR 26-JUN-2001; 2001US-00891609.  
XX  
XX (AARO-) AARON DIAMOND AIDS RES CENT.  
PA  
PI Stamatos L, Barnett S, Shrivastava I;  
XX  
XX WPI: 2002-130836/17.  
DR N-PSDB; AAS15499.  
XX

Immunizing an animal and eliciting an immune response against  
PT heterologous HIV-1 in an animal, involves administering an immunogen  
PT comprising modified HIV-1 envelope protein, or DNA or virus encoding the  
PT protein.  
XX  
XX Claim 5; Fig 17; 62pp; English.  
PS

CC The present invention relates to methods for immunising an animal,  
CC preferably human, against a heterologous human immunodeficiency virus type  
CC 1 (HIV-1), and eliciting a heterologous immune response to HIV-1 in that  
CC animal. The method comprises administering an immunogen having at least  
CC one modified HIV-1 envelope protein with a deletion in the second  
CC hypervariable (V2) region or its fragment, or DNA or virus encoding the  
CC modified HIV-1 envelope protein. The invention discloses the  
CC polynucleotide and amino acid sequences for the full-length HIV-1  
CC SF162deltaV2 (SF162DV2) gp140 envelope protein, and, for an N-terminal  
CC mutant of SF162DV2. The immunogens of the invention are useful for  
CC immunising an animal against heterologous HIV-1 strains by eliciting  
CC neutralising antibodies or protective antibodies in the animal. The  
CC method can be used to treat acquired immunodeficiency syndrome (AIDS).  
CC The present sequence represents the N-terminal mutant (deletion of amino  
CC acids 1-27) of modified HIV-1 SF162DV2 gp140 envelope protein  
XX  
SQ Sequence 619 AA;  
Query Match 95.8%; Score 3327; DB 5; Length 619;  
Best Local Similarity 100.0%; Pred. No. 5.8e-162;  
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 28 SAVEKGLWTVVYGVVPWKEATTLFCASDAKAYDTEVHNWVWATHACVPTDNPQEVLEN 87  
Db 1 SAVEKGLWTVVYGVVPWKEATTLFCASDAKAYDTEVHNWVWATHACVPTDNPQEVLEN 60  
QY 88 VTENFNMMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 147  
Db 61 VTENFNMMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 120  
QY 148 DRGEIKNCSFKVGAGKLNCSNTSVITQACPKVSFPIPIHYCAPAGFAILKCNKDKFNCS 207  
Db 121 DRGEIKNCSFKVGAGKLNCSNTSVITQACPKVSFPIPIHYCAPAGFAILKCNKDKFNCS 180  
QY 208 GECTNVSTVQCTHGIRPVVSTOLLNGSLAEGVVRSENFDTNACTIIVOLKESVEINC 267  
Db 181 GECTNVSTVQCTHGIRPVVSTOLLNGSLAEGVVRSENFDTNACTIIVOLKESVEINC 240  
QY 268 TRPNNTRKSTIGRGRAYATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLQAFGNKT 327  
Db 241 TRPNNTRKSTIGRGRAYATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLQAFGNKT 300  
QY 328 IVFKOSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNTTIGPNNNTGTITLPCRIKQII 387  
Db 301 IVFKOSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNTTIGPNNNTGTITLPCRIKQII 360  
QY 388 NRWQEVGKAMYAPPVIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSEL 447  
Db 361 NRWQEVGKAMYAPPVIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSEL 420  
QY 448 YKYKVVKIEPLGVAPTAKARRVQREKRAVTLGAMFLGAGSTMGARSITLTVQARQ 507  
Db 421 YKYKVVKIEPLGVAPTAKARRVQREKRAVTLGAMFLGAGSTMGARSITLTVQARQ 480  
QY 508 LLSGIVQOONLLRAIEAQHLLQITVMGK-LQARVLAVERYLKDQQLGWTGCGSKLCT 567  
Db 481 LLSGIVQOONLLRAIEAQHLLQITVMGK-LQARVLAVERYLKDQQLGWTGCGSKLCT 540  
QY 568 TTAVPWNASWSNKSLLDQIWNMTWMEWEREIDNYTNLIYTLIEESQOQEKNEQELLELD 627  
Db 541 TTAVPWNASWSNKSLLDQIWNMTWMEWEREIDNYTNLIYTLIEESQOQEKNEQELLELD 600  
QY 628 KWASLWNWFDISKWLWYIK 646  
Db 601 KWASLWNWFDISKWLWYIK 619  
RESULT 4  
ABR55684  
ID ABR55684 standard; protein; 842 AA.  
XX  
XX ABR55684;  
XX

DT 23-OCT-2003 (revised)  
DT 18-AUG-2003 (first entry)  
XX HIV isolate SF162 env polypeptide.  
XX HIV; gag; nef; prot; tat; rev; vif; vpr; vpu; env; anti-HIV; vaccine; immune response.  
XX Human immunodeficiency virus 1.  
XX WO2003020876-A2.  
XX 13-MAR-2003.  
XX 05-JUL-2002; 2002WO-US021342.  
XX 31-AUG-2001; 2001US-0316860P.  
XX 16-JAN-2002; 2002US-0349728P.  
XX (CHIR ) CHIRON CORP.  
XX Zur Megede J, Barnett SW, Lian Y, WPI; 2003-278761/27.  
XX New expression cassettes and polynucleotides encoding HIV Gag, Nef, Prot, Tat, Rev, Vif, Vpr, Vpu, or Env polypeptides, useful for DNA immunization or generating an immune response against HIV in a subject.  
XX Example; Fig 2A-C; 214pp; English.  
XX The invention relates to an expression cassette comprising a polynucleotide sequence encoding a polypeptide including an HIV Gag, Nef, Prot, Tat, Rev, Vif, Vpr, Vpu, or Env polypeptide. The expression cassettes, HIV polypeptides and polynucleotides encoding the HIV polypeptides are useful for DNA immunization or generating an immune response against HIV in a subject. The polynucleotides are also useful for generating packaging cell lines or producing the HIV polypeptides. CC Sequences ABR5684-688 represent env polypeptides from various HIV CC isolates. (Updated on 23-OCT-2003 to standardise OS field)  
XX Sequence 842 AA;  
SQ  
Query Match 94.4%; Score 3279.5; DB 6; Length 842;  
Best Local Similarity 94.2%; Pred. No. 1.9e-159;  
Matches 619; Conservative 2; Mismatches 7; Indels 29; Gaps 3;  
QY 19 LLLGLMLIC-SAVEKLVTVYVYGVVWKEATTTLCFASDAKAYDTEVHNVWATHACVPTD 77  
DB 13 LLCGAVFVSFAVEKLVTVYVYGVVWKEATTTLCFASDAKAYDTEVHNVWATHACVPTD 72  
QY 78 PNPQEVLENVTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLKNAT 137  
DB 73 PNPQEVLENVTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLKNAT 132  
QY 138 NTKSSNWKMDRGEIKNCSPKVGAG-----KLNKNTS 170  
DB 133 NTKSSNWKMDRGEIKNCSPKVTTSIRNKQKEYALFYKLDVVPIDNDNTSKLNKNTS 192  
QY 171 VITOACPKVSFEPIPIHYCAPAGFALLKCNDDKFKNGSGCTNVSTVQCTGIRPVVSTQL 230  
DB 193 VITOACPKVSFEPIPIHYCAPAGFALLKCNDDKFKNGSGCTNVSTVQCTGIRPVVSTQL 252  
QY 231 LINGSIAESGVIRSFNTDNTAKTIIVQLKESVEINCTRPNNTRKSIITIGRPFYATG 290  
DB 253 LINGSIAESGVIRSFNTDNTAKTIIVQLKESVEINCTRPNNTRKSIITIGRPFYATG 312  
QY 291 DIIGDTRQAHCHNSGKWNNTLKQIVTKLAQFGNKTIVFKOSSGGDPEIVHVSFNCGE 350  
DB 313 DIIGDTRQAHCHNSGKWNNTLKQIVTKLAQFGNKTIVFKOSSGGDPEIVHVSFNCGE 372  
QY 351 FFYCNSTQLFNSWTNNITGPNNTNGTITLPCRIKQIINRWQEVGKAMYAPPPIRGQIRCSS 410

DB 373 FFYCNSTQLFNSWTNNITGPNNTNGTITLPCRIKQIINRWQEVGKAMYAPPPIRGQIRCSS 432  
QY 411 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTKARRVV 470  
DB 433 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTKARRVV 492  
QY 471 QREKRAVTGLAMFLGFLGAAGSTMGARSILTLTVQARQLLSGTVQQNNLLRAIEAQHLL 530  
DB 493 QREKRAVTGLAMFLGFLGAAGSTMGARSILTLTVQARQLLSGTVQQNNLLRAIEAQHLL 552  
QY 531 QLTWVGIIK-LOARVLAVERYLKDOQLLGWGGSGKLICTTAVPWNASWNSKSLDQIWNM 589  
DB 553 QLTWVGIIKLOARVLAVERYLKDOQLLGWGGSGKLICTTAVPWNASWNSKSLDQIWNM 612  
QY 590 TWMEWEREIDNYTLIYTLIBESQKNEQELLELDKWSLWVWFDISKWLWYIK 646  
DB 613 TWMEWEREIDNYTLIYTLIBESQKNEQELLELDKWSLWVWFDISKWLWYIK 669  
RESULT 5  
ADCI3218  
ID ADCI3218 standard; protein; 842 AA.  
XX  
AC ADCI3218;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Protein of HIV Type C, SF162 SEQ ID NO 2.  
XX  
KW expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;  
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV.  
XX  
OS Human immunodeficiency virus.  
XX  
PN WO2003004620-A2.  
XX  
PD 16-JAN-2003.  
XX  
PF 05-JUL-2002; 2002WO-US021420.  
XX  
PR 05-JUL-2001; 2001US-0303192P.  
PR 31-AUG-2001; 2001US-0316860P.  
PR 16-JAN-2002; 2002US-0349871P.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (UYST-) UNIV STELLENBOSCH.  
XX  
PI Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EU;  
XX  
XX WPI; 2003-221593/21.  
XX  
PT New expression cassette comprising a polynucleotide sequence encoding a polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide, useful for immunization, or generating packaging cell lines.  
XX  
PS Disclosure; Fig 2; 301pp; English.  
XX  
CC The invention relates to a novel expression cassette comprising a polynucleotide sequence encoding a polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel expression cassette can be used to treat HIV type C by gene therapy or used in the development of a vaccine. The gene delivery vector is administered intramuscularly, intramusosally, intranasally, subcutaneously, intradermally, transdermally, intravaginally, intrarectally, orally or intravenously. The expression cassette is useful for immunisation, generating packaging cell lines and producing HIV CC polypeptides. This sequence represents an HIV Type C related protein of the invention.  
XX  
SQ Sequence 842 AA;  
Query Match 94.4%; Score 3279.5; DB 7; Length 842;

Best Local Similarity 94.2%; Pred. No. 1.9e-159;		Matches 619; Conservative 2; Mismatches 7; Indels 29; Gaps 3;	
QY	19	LLGLMLIC-SAVEKLVVTVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTD	77
DB	13	LLCGAVFVSPSAVEKLVVTVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTD	72
QY	78	PNPQEVLENVTENFMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNT	137
DB	73	PNPQEVLENVTENFMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNT	132
QY	138	NTKSSNWKEMDGEIKNCSFKVAG-----KLINCNTS	170
DB	133	NTKSSNWKEMDGEIKNCSFKVTSIRNMQKEYALFYKLDVVPIDNDNTSYKLINCNTS	192
QY	171	VITQACPVSFPEPIHYCAPAGFAILKCNCKKFGSGPCTNVSTVQCTHGRPVVSTOL	230
DB	193	VITQACPVSFPEPIHYCAPAGFAILKCNCKKFGSGPCTNVSTVQCTHGRPVVSTOL	252
QY	231	LLNGSLAEEGVVIRSENFDTNACTIIVQLKESVEINCTRPNNTRKSIITIGGRAFYATG	290
DB	253	LLNGSLAEEGVVIRSENFDTNACTIIVQLKESVEINCTRPNNTRKSIITIGGRAFYATG	312
QY	291	DIIGDIRQAHCNISGEKWNNTLKQIVTKLOAQFGNKTIVFKOSSGDDPEIVMHSFNCGE	350
DB	313	DIIGDIRQAHCNISGEKWNNTLKQIVTKLOAQFGNKTIVFKOSSGDDPEIVMHSFNCGE	372
QY	351	FFYCNSTOLFNSWTNNTIGPNTNGTITLPCRIKOIINRWQEVGKAMYPPIRGQIRCSS	410
DB	373	FFYCNSTOLFNSWTNNTIGPNTNGTITLPCRIKOIINRWQEVGKAMYPPIRGQIRCSS	432
QY	411	NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYVKVIEPLGVAPTKARRV	470
DB	433	NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYVKVIEPLGVAPTKARRV	492
QY	471	QREKAVTLGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQOQNLLRAIEAQOHL	530
DB	493	QREKAVTLGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQOQNLLRAIEAQOHL	552
QY	531	QUTVWGIK-LOARVLAVERYLKDQQLLGWCGSGKLICTTAVPMNASWSKSLDQIWNM	589
DB	553	QUTVWGIKLOARVLAVERYLKDQQLLGWCGSGKLICTTAVPMNASWSKSLDQIWNM	612
QY	590	TWMEWEREDNTNLIYTLIEESQOEKNEQELLELDKWSLWNWFDISKWLWYIK	646
DB	613	TWMEWEREDNTNLIYTLIEESQOEKNEQELLELDKWSLWNWFDISKWLWYIK	669
RESULT 6			
ID	ABU66565	ABU66565 standard; protein; 842 AA.	
XX	AC	ABU66565;	
XX	DT	22-MAY-2003 (first entry)	
XX	DE	Human immunodeficiency virus (HIV) envelope (env) protein #1.	
XX	KW	Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;	
XX	KW	gene therapy; packaging cell line; humoral immune response;	
XX	KW	cellular immune response; gene delivery vector; DNA immunisation;	
XX	OS	envelope protein; env.	
XX	OS	Human immunodeficiency virus.	
PN	WO2003004657-A1.		
XX	PD	16-JAN-2003.	
XX	PF	05-JUL-2002; 2002WO-US021421.	
XX	PR	05-JUL-2001; 2001US-0303192P.	
PR	31-AUG-2001; 2001US-0316860P.		

PR	16-JAN-2002; 2002US-0349728P.	Query Match Best Local Similarity 94.3%; Score 3274.5; DB 6; Length 842; Matches 618; Conservative 2; Mismatches 8; Indels 29; Gaps 3;	SQ	Sequence 842 AA;
PR	16-JAN-2002; 2002US-034973P.			
PR	16-JAN-2002; 2002US-0349871P.			
XX	(CHIR ) CHIRON CORP.			
PI	Zur Megede J, Barnett SW, Lian Y;	The invention describes a synthetic polynucleotide encoding 2 or more immunogenic HIV polypeptides, where at least 2 of the polypeptides are derived from different HIV subtypes. The polynucleotide is useful for immunisation, generation of packaging cell lines, or production of HIV polypeptides. The polynucleotide and its encoded proteins are useful as immunogenic compositions or vaccines for generating humoral or cellular immune responses against HIV in a subject, or for inducing neutralising antibodies against HIV. The gene delivery vector comprising the polynucleotide is also useful for DNA immunisation of, or for generating an immune response (e.g. a humoral or cellular immune response) in, a subject such as a mammal, particularly a human. This is the amino acid sequence of a human immunodeficiency virus (HIV) envelope (env) protein	CC	Sequence 842 AA;
PI	WPI; 2003-221602/21.			
XX				
XX				
PT	New synthetic polynucleotides encoding antigenic HIV type B and/or type C polypeptides, useful as immunogenic compositions or vaccines for generating humoral or cellular immune responses against HIV in a subject, especially humans.	Disclosure; Fig 2A-C; 262pp; English.	CC	Sequence 842 AA;
PT				
PT				
PT				
PS		The invention describes a synthetic polynucleotide encoding 2 or more immunogenic HIV polypeptides, where at least 2 of the polypeptides are derived from different HIV subtypes. The polynucleotide is useful for immunisation, generation of packaging cell lines, or production of HIV polypeptides. The polynucleotide and its encoded proteins are useful as immunogenic compositions or vaccines for generating humoral or cellular immune responses against HIV in a subject, or for inducing neutralising antibodies against HIV. The gene delivery vector comprising the polynucleotide is also useful for DNA immunisation of, or for generating an immune response (e.g. a humoral or cellular immune response) in, a subject such as a mammal, particularly a human. This is the amino acid sequence of a human immunodeficiency virus (HIV) envelope (env) protein	CC	Sequence 842 AA;
PS				
PS				
PS				
CC		The invention describes a synthetic polynucleotide encoding 2 or more immunogenic HIV polypeptides, where at least 2 of the polypeptides are derived from different HIV subtypes. The polynucleotide is useful for immunisation, generation of packaging cell lines, or production of HIV polypeptides. The polynucleotide and its encoded proteins are useful as immunogenic compositions or vaccines for generating humoral or cellular immune responses against HIV in a subject, or for inducing neutralising antibodies against HIV. The gene delivery vector comprising the polynucleotide is also useful for DNA immunisation of, or for generating an immune response (e.g. a humoral or cellular immune response) in, a subject such as a mammal, particularly a human. This is the amino acid sequence of a human immunodeficiency virus (HIV) envelope (env) protein	CC	Sequence 842 AA;
CC				
CC				
CC				
QY	19	LLGLMLIC-SAVEKLVVTVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTD	77	QY
DB	13	LLCGAVFVSPSAVEKLVVTVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTD	72	
QY	78	PNPQEVLENVTENFMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNT	137	
DB	73	PNPQEVLENVTENFMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNT	132	
QY	138	NTKSSNWKEMDGEIKNCSFKVAG-----KLINCNTS	170	QY
DB	133	NTKSSNWKEMDGEIKNCSFKVTSIRNMQKEYALFYKLDVVPIDNDNTSYKLINCNTS	192	
QY	171	VITQACPVSFPEPIHYCAPAGFAILKCNCKKFGSGPCTNVSTVQCTHGRPVVSTOL	230	
DB	193	VITQACPVSFPEPIHYCAPAGFAILKCNCKKFGSGPCTNVSTVQCTHGRPVVSTOL	252	
QY	231	LLNGSLAEEGVVIRSENFDTNACTIIVQLKESVEINCTRPNNTRKSIITIGGRAFYATG	290	QY
DB	253	LLNGSLAEEGVVIRSENFDTNACTIIVQLKESVEINCTRPNNTRKSIITIGGRAFYATG	312	
QY	291	DIIGDIRQAHCNISGEKWNNTLKQIVTKLOAQFGNKTIVFKOSSGDDPEIVMHSFNCGE	350	
DB	313	DIIGDIRQAHCNISGEKWNNTLKQIVTKLOAQFGNKTIVFKOSSGDDPEIVMHSFNCGE	372	
QY	351	FFYCNSTOLFNSWTNNTIGPNTNGTITLPCRIKOIINRWQEVGKAMYPPIRGQIRCSS	410	QY
DB	373	FFYCNSTOLFNSWTNNTIGPNTNGTITLPCRIKOIINRWQEVGKAMYPPIRGQIRCSS	432	
QY	411	NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYVKVIEPLGVAPTKARRV	470	
DB	433	NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYVKVIEPLGVAPTKARRV	492	
QY	471	QREKAVTLGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQOQNLLRAIEAQOHL	530	QY
DB	493	QREKAVTLGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQOQNLLRAIEAQOHL	552	
QY	531	QUTVWGIK-LOARVLAVERYLKDQQLLGWCGSGKLICTTAVPMNASWSKSLDQIWNM	589	

Db 553 QLTVMGIKQARVLAVERYLKDQLLGWCSGLICTTAVPWNASWNSKSLDQIWNMM 612

Qy 590 TWMEWERIDNTNLIYTLIIESNQOKEQKELELDKWSLWNPFDISKWLIWYIK 646

Db 613 TWMEWERIDNTNLIYTLIIESNQOKEQKELELDKWSLWNPFDISKWLIWYIK 669

RESULT 7	
ABB06211	
ID	ABB06211 standard; protein; 842 AA.
XX	
AC	ABB06211;
XX	
DT	15-MAY-2002 (first entry)
XX	
DE	HIV Env isolate SF162 amino acid sequence.
XX	
KW	Human immunodeficiency virus type C; antigenic HIV type C protein;
KW	immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW	immunostimulant; gene therapy.
XX	
OS	Human immunodeficiency virus; type C.
OS	Synthetic.
XX	
PN	W0200204493-A2.
XX	
PD	17-JAN-2002.
XX	
PF	05-JUL-2001; 2001WO-US021241.
XX	
PR	05-JUL-2000; 2000US-00610313.
XX	
PA	(CHIR ) CHIRON CORP.
PA	(UYST-) UNIV STELLENBOSCH.
XX	
PI	Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX	
DR	WPI; 2002-154920/20.
XX	
PT	New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT	applications including DNA immunisation or generation of packaging cell
PT	lines, particularly in gene therapy.
XX	
PS	Disclosure; Fig 105; 233pp; English.
XX	
CC	The present invention describes expression cassettes comprising a
CC	polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC	type C polypeptides. The expression cassettes comprise any of the HIV
CC	type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC	(1). (I) have immunostimulant activity and can be used in gene therapy.
CC	The HIV type C polynucleotides are useful in applications including DNA
CC	immunisation, generation of packaging cell lines, and production of HIV
CC	Type C proteins. The polynucleotides are particularly useful in gene
CC	therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC	ABB06204 to ABB06215 represent sequences used in the exemplification of
XX	the present invention
XX	
SQ	Sequence 842 AA;

Query Match	94.0%;	Score	3264.5;	DB	5;	Length	842;
Best Local Similarity	93.8%;	Pred. No.	1.1e-158;				
Matches	616;	Conservative	2;	Mismatches	10;	Indels	29;
Gaps	3						
Qy	19	LLGLMLMIC-SAVEK	LWVTVYVYGV	PWKAEATTLF	CASDAKAYDTEV	HNWATHACVPTD	77
			:				
			:				
Db	13	LLCGAVFVSPSAVEK	LWVTVYVYGV	PWKAEATTLF	CASDAKAYDTEV	HNWATHACVPTD	72
Qy	78	PNPQEVILENVTFENF	NKNNVVEQW	HEDIISLMDQSLKPCV	KLTPLCVTLHCTNLK	KNAT	137
Db	73	PNPQEVILENVTFENF	NKNNVVEQW	HEDIISLMDQSLKPCV	KLTPLCVTLHCTNLK	KNAT	132
Qy	138	NTKSSNWKMDRGEIK	NCNCSFKVGAG	-----	-----	KLINGNTS	170

Db	133	NTKSSNWEMDRGEIKNCSEFKVTTTSIRNKMQKEYALFYKLDVVPIQNDNTSYKLIINCNTS	192
Qy	171	VITQACPQVSEPIPIHYHCAPAGPAILKCNCKFNKFGPCTNVSTVQCTHGIRPVVSTQL	230
Db	193	VTTQACPQVSEPIPIHYHCAPAGPAILKCNCKFNKFGPCTNVSTVQCTHGIRPVVSTQL	252
Qy	231	LLNGSLABEGVVIRSEFTDNAKTIIVOLKESVEINCTRNNTNRKSTITIGPGRAFAYTG	290
Db	253	LLNGSLABEGVVIRSEFTDNAKTIIVOLKESVEINCTRNNTNRKSTITIGPGRAFAYTG	312
Qy	291	DIIGDIROAHCHNISGEKWNNTLKOIVTKLOAQFGNKTIIVFKQSSGGDPEIWMHSEFNCGGE	350
Db	313	DIIGDIROAHCHNISGEKWNNTLKOIVTKLOAQFGNKTIIVFKQSSGGDPEIWMHSEFNCGGE	372
Qy	351	PFYCNSTOLFNSWTNNTTIGPNTNGTITLPCRKQIINRWQEVGKAMYPPIRGQIRCSS	410
Db	373	PFYCNSTOLFNSWTNNTTIGPNTNGTITLPCRKQIINRWQEVGKAMYPPIRGQIRCSS	432
Qy	411	NITGLLLTRDGGKEISNTTEIFRPGGDMRDNMRSELYKYKVKVIEPLGVAPTKAKRRVV	470
Db	433	NITGLLLTRDGGKEISNTTEIFRPGGDMRDNMRSELYKYKVKVIEPLGVAPTKAKRRVV	492
Qy	471	QREKRAVTLGAMFLGFLGAAGSTWGARSLLTITVQARQLLSGIVQQQNNLLRAIEAQOHL	530
Db	493	QREKRAVTLGAMFLGFLGAAGSTWGARSLLTITVQARQLLSGIVQQQNNLLRAIEAQOHL	552
Qy	531	QLTWVGIK-LOARVLAVERYLKDOQLLGICGCSGLICTTAVPNASWSNKSLLDQIWNMM	589
Db	553	QUTWVGIIKQLQARVLAVERYLKDOQLLGICGCSGLICTTAVPNASWSNKSLLDQIWNMM	612
Qy	590	TWMEWERIDNVTNLIYTLIEESNQEFKNEQELLELDKWASLWNWFDISKWLWYIK	646
Db	613	TWMEWERIDNVTNLIYTLIEESNQEFKNEQELLELDKWASLWNWFDISKWLWYIK	669
RESULT 8			
AAB82761			
ID	AAB82761 standard; protein; 883 AA.		
AC	XX		
CC	AAB82761;		
XX	XX		
DT	11-SEP-2003 (revised)		
DT	29-OCT-2001 (first entry)		
XX	XX		
DE	Ancestral HIV-1 group M, subtype B gp160 protein.		
XX	XX		
KW	HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.		
OS	Human immunodeficiency virus 1.		
XX	XX		
PN	W0200160838-A2.		
XX	XX		
PD	23-AUG-2001.		
XX	XX		
PF	16-FEB-2001; 2001WO-US005288.		
XX	XX		
PR	18-FEB-2000; 2000US-0183659P.		
XX	XX		
PA	(UNIW ) UNIV WASHINGTON.		
XX	XX		
PI	Mullins JI, Rodrigo AG, Learn GH, Li F;		
XX	XX		
DR	WPI; 2001-536565/59.		
XX	N-PSDB; AAB26468.		
XX	XX		
PT	Preparing an ancestral viral amino acid sequence useful as a vaccine		
PT	comprise determining a recent ancestor of a circulating virus by maximum		
PT	likelihood phylogeny analysis.		
XX	XX		
PS	Claim 8; Page 54; 89pp; English.		
XX	XX		
CC	The present sequence is that of an ancestral HIV-1 group M subtype B		

CC gp160 (env gene product) sequence. The invention provides compositions  
CC and methods for determining ancestral viral gene sequences and ancestral  
CC viral protein sequences for highly diverse viruses, such as HIV-1. The  
CC methods use samples of circulating viruses to determine an ancestral  
CC viral sequence by maximum likelihood phylogeny analysis. In the present  
CC case, the ancestral HIV-1 subtype B env sequence (see AAH26468) was  
CC determined using 38 subtype B sequences (obtained from 9 different  
CC countries) and 3 subtype D (outgroup) sequences. The distances between  
CC this ancestral viral sequence and circulating strains used to determine  
CC it were on average 12.3% (range: 8.0-21.0%) while the available specimens  
CC were 17.3% different from each other (range: 13.3-23.2%). Thus, the  
CC ancestor sequence was, on average, more closely related to any given  
CC circulating virus than to any other variant. The ancestral gp160 sequence  
CC included a wide variety of immunogenically active peptides when processed  
CC for antigen presentation; nearly all known subtype B CTL epitope  
CC consensus amino acids were represented. Thus, an immunogenic composition  
CC to this subtype B ancestor protein will elicit broad neutralising  
CC antibody against HIV-1 isolates of the same subtype, and will also elicit  
CC a broad cellular response mediated by antigen-specific T-cells. A claimed  
CC vaccine composition comprises a viral ancestor protein or its immunogenic  
CC fragment, especially one derived from the HIV-1 group M subtype B gp160  
CC ancestral protein. (Updated on 11-SEP-2003 to standardise OS field)  
XX  
XX Sequence 883 AA;  
SQ

Query Match 88.2%; Score 3065; DB 4; Length 883;  
Best Local Similarity 83.7%; Pred. No. 1.7e-148;  
Matches 594; Conservative 17; Mismatches 35; Indels 64; Gaps 10;

QY 1 MRVKGIRKXNYOHLWRGGTLLGLMLCSAVEKLWTVVYGVVPVWKEATTLFCASDAKAY 60  
DB 1 MRVKGIRKXNYOHLWRGGTLLGLMLCSAAEKLWTVVYGVVPVWKEATTLFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDPNPQBIENVNTENFMNKNMVEQMHEDIISLWQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDPNPQBIENVNTENFMNKNMVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPCLCVTLHCTN--LKNATNTKSN-----WKEMDRGEIKNCSFKVGAG----- 162  
DB 121 LTPCLCVTLNCTDLRLNATNTNSSATNTTSSGGTMEGKEIKNCSFNVTTSDRM 180  
QY 163 -----KLNCNTSVITQACPVSFEPPIHYCAPAGF 194  
DB 181 QKEYALFYKLDVVPIDNNNTNNTSYRLNCNTSVITQACPVSFEPPIHYCTTAPG 240  
QY 195 AILKNDKFKNGSGPCTNVSTVQCTGIRPVVSTOLLNGSLAEQGVIRSNFTDNAKT 254  
DB 241 AILKNDKFKNGSGPCTNVSTVQCTGIRPVVSTOLLNGSLAEQGVIRSNFTDNAKT 300  
QY 255 IIVQLKESVEINCTRNNTNRKSIITIGPGRAFYATGDIIGDIRQAFNCISGKWNNTLKQ 314  
DB 301 IIVQLNESVEINCTRNNTNRKSIPIGPGRALYATGKIIGDIRQAHCNLSRAKWNNTLKQ 360  
QY 315 IYTKLOAQGNK--TIYFKQSGGDEPVIWHSFNGCGEFPYCNSTQLFNSTW--NNTIGP 370  
DB 361 IYTKLREQFGNNKTTIYFNQSGGDEPVIWHSFNGCGEFPYCNSTQLFNSTWTFNGTGN 420  
QY 371 NNT-----NGTITLPCRKIQIINRWQEVGKAWYAPPRIQRCSSNITGLLLTRDG 421  
DB 421 NNTERSNNAADDNTITLPCRKIQIINRWQEVGKAWYAPPISQIRCSSNITGLLLTRDG 480  
QY 422 G--KEISNT--TBIFRPGGDMRDNRSLEYKYKVKVIEPLGVAPTAKERVVQREKRAY- 477  
DB 481 GNNENTNTDTETFRPGGDMRDNRSLEYKYKVKVIEPLGVAPTAKERVVQREKRAYG 540  
QY 478 TIGAMPLFLGAAGSTMGARSUTLTQVQRLISGIVQOQNLLRAIEAQHLLQLTVWGI 537  
DB 541 MLGAMPLFLGAAGSTMGASMTLTQVQRLISGIVQOQNLLRAIEAQHLLQLTVWGI 600  
QY 538 K-LQARVLAVERYLKDQQLLWGCSGKLICTTAVPWNASWNSKSDQIWNNTWMEWER 596  
DB 601 KQLOARVLAVERYLKDQQLLWGCSGKLICTTAVPWNASWNSKSDQIWNNTWMEWER 660

QY 597 EIDNTNLIYTIERSQOQEKNEQELLELDKWSLNNWFDISKMLWYIK 646  
DB 661 EIDNTYGLIYTIERSQOQEKNEQELLELDKWSLNNWFDITNMLWYIK 710

RESULT 9  
AAR67724  
ID AAR67724 standard; protein; 850 AA.  
XX AAR67724;  
XX  
DT 16-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 07-SEP-1995 (first entry)  
XX  
gp120 from the HIV GNE8 isolate.  
XX  
KW HIV; human immunodeficiency virus; gp120; glycoprotein; GNE clone;  
KW GNE8 isolate.  
XX  
OS Human immunodeficiency virus 1.  
XX  
PN WO9428929-A1.  
XX  
PD 22-DEC-1994.  
XX  
PF 07-JUN-1994; 94WO-US006036.  
XX  
PR 07-JUN-1993; 93US-00072833.  
XX  
PA (GETH ) GENENTECH INC.  
PI Berman PW, Nakamura GR;  
XX  
XX WPI: 1995-036112/05.  
DR N-PSDB; AAQ76018.  
XX  
PT Use of HIV gp 120 polypeptide(s) - for developing probes for the  
analysis, prevention and therapy of HIV infection.  
XX  
PS Claim 23; Page 27-30; 108pp; English.  
XX  
CC This protein is the gp120 (envelope glycoprotein) from the GNE8 isolate  
of HIV-1 (human immunodeficiency virus type 1). There are neutralising  
epitopes in the V2 and C4 domains of gp120, in addition to the  
neutralising epitopes in the V3 domain. Although the amino acid sequences  
of the neutralising epitopes in the V2, V4 and C4 domains are variable,  
the amount of variation is highly constrained. This facilitates the  
design of HIV subunit vaccines that can induce antibodies that neutralise  
the most common HIV strains for a given geographic region. This invention  
provides a multivalent gp120 subunit vaccine where the gp120 present in  
the vaccine is from at least two HIV isolates which have different amino  
acid sequences for a neutralising epitope in these regions. (Updated on  
25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise  
OS field)  
XX  
SQ Sequence 850 AA;  
Query Match 88.2%; Score 3062.5; DB 2; Length 850;  
Best Local Similarity 85.3%; Pred. No. 2.2e-148;  
Matches 580; Conservative 28; Mismatches 35; Indels 37; Gaps 5;

QY 1 MRVKGIRKXNYOHLWRGGTLLGLMLCSAVEKLWTVVYGVVPVWKEATTLFCASDAKAY 60  
DB 1 MRVKGIRKXNYOHLWRGGTLLGLMLCSAAEKLWTVVYGVVPVWKEATTLFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDPNPQBIENVNTENFMNKNMVEQMHEDIISLWQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDPNPQBIENVNTENFMNKNMVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPCLCVTLHCTNLTNATNTKSNWKEMDRGEIKNCSFKVGAG----- 162  
DB 121 LTPCLCVTLNCTDLKNTNTTSSWGMEGKEIKNCSFNVTTSDRMKNNEYALFYKLDV 180





RESULT 11  
 AAP80967  
 ID AAP80967 standard; protein; 851 AA.  
 AC AAP80967;  
 XX  
 DT 24-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 15-NOV-1990 (first entry)  
 XX  
 DE HIV protein HT6.  
 XX  
 KW HIV; HT6; gp160; envelope protein; RF; AIDS.  
 XX  
 OS Human immunodeficiency virus; variant RF.  
 XX  
 PN EP272858-A.  
 XX  
 PD 29-JUN-1988.  
 XX  
 PF 14-DEC-1987; 87EP-00310967.  
 XX  
 PR 15-DEC-1986; 86US-00941111.  
 PR 31-AUG-1987; 87US-00091481.  
 XX  
 PA (REPK ) REPLIGEN CORP.  
 XX  
 PI Rusche J, Lynn D, Carson H, Putney S, Jellis CL;  
 XX  
 DR WPI; 1988-176944/26.  
 DR N-PSDB; AAN80948.  
 XX  
 PT Prodn. of recombinant HIV envelope proteins in insect cells - useful as  
 XX vaccine against AIDS and for diagnosis and therapy.  
 PS Disclosure; Page ?; 4pp; English.  
 XX  
 CC The sequence is the result of cloning a hybrid envelope gene from HIV  
 CC variants BH10 and RF. A central portion of the RF gene was used, the rest  
 CC being from te distantly related variant BH10. The resulting clone, pACHT6  
 CC produces a hybrid gp 160 envelope protein with novel immunological and  
 CC antigenic characteristics. It may be used to as a vaccine and for  
 CC diagnosis and therapy of AIDS. See also AAP80966. (Updated on 25-MAR-2003  
 CC to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 851 AA;  
 Query Match 86.3%; Score 2997; DB 1; Length 851;  
 Best Local Similarity 83.0%; Pred. No. 4.9e-145;  
 Matches 565; Conservative 40; Mismatches 38; Indels 38; Gaps 7;  
 QY 1 MRVKGIRKXQYHL-----WGGTLLGLNLMICSAVEKLVTVYGVVWKEATITLFCASD 56  
 DB 1 MRVK-----EKYQHLWRGWRGWTMLLMGLMICSAVEKLVTVYGVVWKEATITLFCASD 57  
 QY 57 AKAYDTEVHVWATHACVTPDPNPQEI VLENTVFENFMKNMKNVQOMHEDIISLWDQSLK 116  
 DB 58 AKAYDTEVHVWATHACVTPDPNPQEI VLENTVFENFMKNMKNVQOMHEDIISLWDQSLK 117  
 QY 117 PCVKLTPLCVTLHCTNLKATNTKSSNWE-MDRGEIKNCSFKVGA---GK----- 163  
 DB 118 PCVKLTPLCVSLKCTDLKNDTNTSSSGRMIMEKEGKNCNFMNMSIRGKVQKEYAFY 177  
 QY 164 -----LINCNTSVITQACPKVSEPIPIHYCAPAGPAILKCNKDPNGSG 208  
 DB 178 KLDIIPIDNTTSYTLTSCNTSVITQACPKVSEPIPIHYCAPAGPAILKCNKDPNGSG 237  
 QY 209 PCTNVSTVQCTGIRPVSTOLLNGSLAEEGVVIRSENFNTONAKTIIVOLKESVEINCT 268  
 DB 238 PCTNVSTVQCTGIRPVSTOLLNGSLAEEGVVIRSENFNTONAKTIIVOLNASVQINCT 297

QY 269 RPNNTKRSITIGPGRAFVATGDIIGDIOAHNCISGEKWNNTLKQIVTKLQAFQNKTI 328  
 DB 298 RPNNTKRSITIGPGRFVIYATGQIIIGDIRKAHCNLSRAQWNNTLKQVVTKLREQPNKTI 357  
 QY 329 VFKQSSGGDPEIVMHSFNCGGFFFCVNSTOLFSTWNNTIGNNTNG--TITLPCRKQI 386  
 DB 358 VFTSSGGDPEIVLHSFNCGGFFFCVNTTOLFSTWNSTEGSNNTGNDTITLPCRKQI 417  
 QY 387 INRWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSE 446  
 DB 418 VNMQEVGKAMYAPPIISGQIKCISNITGLLLTRDGGEDTNTTEIFRPGGDMRDNRSE 477  
 QY 447 LYKYKVKIEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAAGSTMGARSLLTLVQAR 506  
 DB 478 LYKYKVKIEPLGVAPTAKRRVQREKRAVGIGALFLGFLGAAGSTMGASMTLTVQAR 537  
 QY 507 QLLSGIVQOONNLLRAIRAIAQOHLQLTVMGIK-LQARVLAVERYLKDQQLGIGWCGSKL 565  
 DB 538 QLLSGIVQOONNLLRAIRAIAQOHLQLTVMGIK-LQARVLAVERYLKDQQLGIGWCGSKL 597  
 QY 566 ICTTAVPWNASWSNKSLEQIMNNMTMWEWERIDNYNLIYTLIEESQOQEKNEQELLE 625  
 DB 598 ICTTAVPWNASWSNKSLEQIMNNMTMWEWDREINNYTSLIHSLEESQOQEKNEQELLE 657  
 QY 626 LDKWASLWNNFDISKWLWYIK 646  
 DB 658 LDKWASLWNNFNTNWLWYIK 678  
 RESULT 12  
 AAW53112  
 ID AAW53112 standard; protein; 855 AA.  
 XX  
 AC AAW53112;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 25-JUN-1998 (first entry)  
 XX  
 DE ENV protein contained in a complete ARV-2 nucleotide sequence.  
 XX  
 KW ARV-2; enhanced promoter; gene expression; cytomegalovirus; HIV; AIDS;  
 KW ENV protein.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN US5688688-A.  
 XX  
 PD 18-NOV-1997.  
 XX  
 PF 10-AUG-1994; 94US-00288336.  
 XX  
 PR 31-OCT-1984; 84US-00667501.  
 PR 30-JAN-1985; 85US-00696534.  
 PR 06-SEP-1985; 85US-00773447.  
 PR 24-DEC-1987; 87US-00138894.  
 PR 17-AUG-1992; 92US-00931191.  
 PR 28-JUN-1993; 93US-00083391.  
 PR 17-AUG-1993; 93US-00107377.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Haigwood NI, Dina D, Rosenberg S, Chapman BS, Thayer RM;  
 PI Luciw PA;  
 XX  
 DR WPI; 1998-007982/01.  
 DR N-PSDB; AAV04733.  
 XX  
 PT Enhanced promoter for gene expression - comprising cytomegalovirus  
 PT immediate early promoter plus intron.  
 XX  
 PS Example 1; Fig 4C-P; 99pp; English.  
 XX  
 CC This sequence represents the ENV protein contained in a complete

CC nucleotide sequence of ARV-2 derived from partial sequences of several  
CC ARV clones. The invention provides a method for construction of a vector  
CC for expression of a polypeptide in a mammalian cell, comprising a  
CC polypeptide coding sequence operably linked downstream of an enhanced  
CC promoter. The enhanced promoter comprises the human cytomegalovirus  
CC immediate early region (HCMV IE1) promoter and the first intron proximate  
CC to the 3' end of the HCMV IE1 promoter. The polypeptide can be any of the  
CC HIV recombinant polypeptides and especially HIV gp120. Expression of HIV  
CC gp120 by COS 7 cells transfected with pCMV6a containing the gp120 coding  
CC region, where pCMV6a is a vector containing the above enhanced promoter,  
CC is increased by a factor of 50-100 compared with the use of a vector  
CC containing the SV40 early promoter. (Updated on 17-OCT-2003 to  
CC standardise OS field)  
XX  
XX Sequence 855 AA;

Query Match 85.9%; Score 2984; DB 2; Length 855;  
Best Local Similarity 82.4%; Pred. No. 2.2e-144;  
Matches 563; Conservative 34; Mismatches 48; Indels 38; Gaps 6;  
QY 1 MRVKGIRKRYQHWRGTTLLGLMLICSAVEKLWTVVYGVVPWKAEATTLFCASDAKAY 60  
DB 1 MKVKGTRRYQHWRGTTLLGLMLICSAVEKLWTVVYGVVPWKAEATTLFCASDARAY 60  
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNMMKNMKNMVEQMHEDIISLWQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNMMKNMKNMVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPLCVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSEFNITTSIRDKIQENALFRNLDDV 162  
DB 121 LTPLCVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSEFNITTSIRDKIQENALFRNLDDV 180  
QY 163 -----KLINCVTSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKNGSG 208  
DB 181 PIDNASTTTNTYNYRLIHCNRSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKNGSG 240  
QY 209 PCTNVSTVQCTHGIRPWSVQTLNLGSLABEGVVRSENFNTONAKTIIIVOLKESVEINCT 268  
DB 241 PCTNVSTVQCTHGIRPWSVQTLNLGSLABEGVVRSENFNTONAKTIIIVOLKESVEINCT 300  
QY 269 RPNNTNRKSIITGPGRAFVATGDIIGDIRQAHNCISGEKWNNTLKOIVTKLQAQFG-NKT 327  
DB 301 RPNNTNRKSIITGPGRAFVATGDIIGDIRQAHNCISGEKWNNTLKOIVTKLQAQFG-NKT 360  
QY 328 IVFNQSSGGDPEIWMHSCNCGEFCSTOLPNSWT--NNTIGPNTNGTITLPCRQK 385  
DB 361 IVFNQSSGGDPEIWMHSCNCGEFCSTOLPNSWT--NNTIGPNTNGTITLPCRQK 419  
QY 386 IINRWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGKISNTTEIFRPGGDMRDNRWS 445  
DB 420 IINRWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGKISNTTEIFRPGGDMRDNRWS 479  
QY 446 ELYKYKVKIIEPLGVAPTKAKRRVQREKRAV--TLGAMFLGFLGAAGSTMGARSFLTIVQ 504  
DB 480 ELYKYKVKIIEPLGVAPTKAKRRVQREKRAV--TLGAMFLGFLGAAGSTMGARSFLTIVQ 539  
QY 505 ARQLLSGIVQOQNLLRATRAQOHLQLTWGK-LQARVLAVERYLKQOQLLIGWCSG 563  
DB 540 ARQLLSGIVQOQNLLRATRAQOHLQLTWGK-LQARVLAVERYLKQOQLLIGWCSG 599  
QY 564 KLCTTAVPNWASWSKSLDOIWNNTWMWEIREIDNTNLIYTLLEESQNOQEKNEQEL 623  
DB 600 KLCTTAVPNWASWSKSLDOIWNNTWMWEIREIDNTNLIYTLLEESQNOQEKNEQEL 659  
QY 624 LELEDKASLWNNFDSIKLWYIK 646  
DB 660 LELEDKASLWNNFDSIKLWYIK 682  
RESULT 13  
AAY77298  
ID AAY77298 standard; protein; 855 AA.  
XX

AC AAY77298;  
XX 12-SEP-2003 (revised)  
DT 22-MAY-2000 (first entry)  
DE HIV-1 (ATCC CRL 8597) env protein.  
XX HIV-1; immunoassay; antigen; pol fragment; p31; recombinant; antibody;  
KW detection; env protein.  
XX Human immunodeficiency virus 1; 'ATCC CRL 8597'.  
OS US6013432-A.  
XX 11-JAN-2000.  
XX 17-MAY-1995; 95US-00443434.  
XX 31-OCT-1984; 84US-00667501.  
PR 30-JAN-1985; 85US-00696534.  
PR 06-SEP-1985; 85US-00773447.  
PR 24-DEC-1987; 87US-00138894.  
PR 17-AUG-1992; 92US-00931154.  
PR 08-JUL-1993; 93US-00089407.  
XX (CHIR ) CHIRON CORP.  
XX Luciw PA, Dina D;  
PI WPI; 2000-170256/15.  
XX N-PSDB; AAZ92021.  
XX Immunoassay for antibodies against human immune deficiency virus, for  
PT diagnosing infection, uses an immunogenic fragment of the pol protein as  
PT antigen.  
XX Example 1; Fig 4K-O; 99pp; English.  
XX The invention relates to the improvement of HIV-1 immunoassays by the use  
CC of an HIV-1 antigen comprising an immunogenic fragment of recombinant or  
CC synthetic HIV-1 pol, which is encoded by an approximately 9.7 kb sequence  
CC between a BstXI restriction site at position 3006 and an NdeI site at  
CC position 5131 of the genome (the proviral DNA sequence is given in  
CC AA920201). The immunogenic pol fragment is not immunologically cross-  
CC reactive with human T cell lymphotropic viruses I or II. The invention  
CC also encompasses the use of p31 as an antigen. The recombinant antigens  
CC may be produced in Escherichia coli, Saccharomyces cerevisiae or in  
CC mammalian cells. Immunoassays using the recombinant HIV proteins may be  
CC used to diagnose and stage HIV-1 infections. Sequences AAY77294-Y77299  
CC represent proteins encoded by the genome of HIV-1 (ATCC CRL 8597).  
XX (Updated on 12-SEP-2003 to standardise OS field)  
XX Sequence 855 AA;  
SQ Query Match 85.9%; Score 2984; DB 3; Length 855;  
Best Local Similarity 82.4%; Pred. No. 2.2e-144;  
Matches 563; Conservative 34; Mismatches 48; Indels 38; Gaps 6;  
QY 1 MRVKGIRKRYQHWRGTTLLGLMLICSAVEKLWTVVYGVVPWKAEATTLFCASDAKAY 60  
DB 1 MKVKGTRRYQHWRGTTLLGLMLICSAVEKLWTVVYGVVPWKAEATTLFCASDARAY 60  
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNMMKNMKNMVEQMHEDIISLWQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNMMKNMKNMVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPLCVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSEFNITTSIRDKIQENALFRNLDDV 162  
DB 121 LTPLCVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSEFNITTSIRDKIQENALFRNLDDV 180  
QY 163 -----KLINCVTSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKNGSG 208  
DB 181 PIDNASTTTNTYNYRLIHCNRSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKNGSG 240







Db 1 MIVGIRKNCQHLWRGTMLLGMLMICSAAEKLWTVVYGVPMKEATTTLFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDNPQIIVLENTENFMKNMVEOMHEDIISLWQSLKPCVK 120  
Db 61 DTEVHNWATHACVPTDNPQIIVLENTENFMKNMVEOMHEDIISLWQSLKPCVK 120  
QY 121 LTPCLVTLHCTNLKNAATNTKSNWKMEDRGEIKNGSFVKYAG-----162  
Db 121 LTPCLVTLNCTDLKNAATNTSSWGKMERGEIKNSFNVTTSIRDKMNEYALFYKLDVV 180  
QY 163 -----KLINCNTSVITQCPKVSPEPIPIHYCAPAGFAILKCNCKKFGSGPCTNV 213  
Db 181 PIDNDNTSVRLISCNSTSVITQCPKVSPEPIPIHYCAPAGFAILKCRDKKFGTGTCTNV 240  
QY 214 STVQCHTGRIPVSVTQLLNGSLAEVGVIRENFTDNAKTIIVOLKESVEINCTRPNN 273  
Db 241 STVQCHTGRIPVSVTQLLNGSLAEVGVIRENFTDNAKTIIVQLNESVEINCTRPNN 300  
QY 274 TRKSITIGPRAFYATGDIIGIROAHCNISGEKWNNTLKQIVTKLQAGFGNKTIVFKOS 333  
Db 301 TRSHIIGPRAFYATGEIIGIROAHCNLSSTKWNNTLKQIVTKLREHF-NKTIVFNHS 359  
QY 334 SGGDPEIVMHSFNCGGEFFYCNSTOLFNS-----TWNTIGPNNNTNGTITLPCRIKQII 387  
Db 360 SGGDPEIVMHSFNCGGEFFYCNSTOLFNSFNVTYTWNTTEGSDTGRNITLQCRKQII 419  
QY 388 NPMQEVGKAMYAPPRIQIRCSNITGLLLTRDGGKEISNTTEIPRPGGDMRDNRSEL 447  
Db 420 NPMQEVGKAMYAPPRIQIRCSNITGLLLTRDGGN--NSETIIPRPGGDMRDNRSEL 477  
QY 448 YKVKVVKIEPLGVAPTKAKRRVQREKRAVTLGAMFLGAGSTMGARSITLTAVQARQ 507  
Db 478 YKVKVVKIEPLGVAPTKAKRRVQREKRAVGAVFLGFLGAGSTMGAASVTLTVQARL 537  
QY 508 LLSGIVQOQNLLRAIEAQOHLQLTVWGIK-LQARVLAVERYLKDQQLLGIWGCSGKLI 566  
Db 538 LLSGIVQOQNLLRAIEAQOHLQLTVWGIKLQARVLAVERYLKDQQLLGIWGCSGKLI 597  
QY 567 CTTAVPWNASWNSKSLDQIWNNTMWEEREIDNVTLSLYIESQOQKNEQELLEL 626  
Db 598 CTTAVPWNASWNSKSLDKIWNNTMWEEREIDNVTLSLYIESQOQKNEQELLEL 657  
QY 627 DKWASLWNNFDISKWLWYIK 646  
Db 658 DKWASLWNNFDTIKWLWYIK 677

RESULT 2

US-09-134-075-28  
; Sequence 28, Application US/09134075  
; Patent No. 6042836  
; GENERAL INFORMATION:  
; APPLICANT: Berman, Phillip W.  
; APPLICANT: Nakamura, Gerald R.  
; TITLE OF INVENTION: HIV Envelope Polypeptides  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP  
; STREET: 3 Embarcadero Center  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/134,075  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/448,603  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haliday, Emily  
; REGISTRATION NUMBER: 38903  
; REFERENCE/DOCKET NUMBER: 14918-704  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-393-2000  
; TELEFAX: 415-393-2286  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 850 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-134-075-28  
  
Query Match 88.2%; Score 3062.5; DB 3; Length 850;  
Best Local Similarity 85.3%; Pred. No. 9,1e-245;  
Matches 580; Conservative 28; Mismatches 35; Indels 37; Gaps 5;  
  
QY 1 MRVKIRKNYQHLWRGTMLLGMLMICSAAVEKLWTVVYGVPMKEATTTLFCASDAKAY 60  
Db 1 MIVGIRKNCQHLWRGTMLLGMLMICSAAEKLWTVVYGVPMKEATTTLFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDNPQIIVLENTENFMKNMVEOMHEDIISLWQSLKPCVK 120  
Db 61 DTEVHNWATHACVPTDNPQIIVLENTENFMKNMVEOMHEDIISLWQSLKPCVK 120  
QY 121 LTPCLVTLHCTNLKNAATNTKSNWKMEDRGEIKNGSFVKYAG-----162  
Db 121 LTPCLVTLNCTDLKNAATNTSSWGKMERGEIKNSFNVTTSIRDKMNEYALFYKLDVV 180  
QY 163 -----KLINCNTSVITQCPKVSPEPIPIHYCAPAGFAILKCNCKKFGSGPCTNV 213  
Db 181 PIDNDNTSVRLISCNSTSVITQCPKVSPEPIPIHYCAPAGFAILKCRDKKFGTGTCTNV 240  
QY 214 STVQCHTGRIPVSVTQLLNGSLAEVGVIRENFTDNAKTIIVOLKESVEINCTRPNN 273  
Db 241 STVQCHTGRIPVSVTQLLNGSLAEVGVIRENFTDNAKTIIVQLNESVEINCTRPNN 300  
QY 274 TRKSITIGPRAFYATGDIIGIROAHCNISGEKWNNTLKQIVTKLQAGFGNKTIVFKOS 333  
Db 301 TRSHIIGPRAFYATGEIIGIROAHCNLSSTKWNNTLKQIVTKLREHF-NKTIVFNHS 359  
QY 334 SGGDPEIVMHSFNCGGEFFYCNSTOLFNS-----TWNTIGPNNNTNGTITLPCRIKQII 387  
Db 360 SGGDPEIVMHSFNCGGEFFYCNSTOLFNSFNVTYTWNTTEGSDTGRNITLQCRKQII 419  
QY 388 NPMQEVGKAMYAPPRIQIRCSNITGLLLTRDGGKEISNTTEIPRPGGDMRDNRSEL 447  
Db 420 NPMQEVGKAMYAPPRIQIRCSNITGLLLTRDGGN--NSETIIPRPGGDMRDNRSEL 477  
QY 448 YKVKVVKIEPLGVAPTKAKRRVQREKRAVTLGAMFLGAGSTMGARSITLTAVQARQ 507  
Db 478 YKVKVVKIEPLGVAPTKAKRRVQREKRAVGAVFLGFLGAGSTMGAASVTLTVQARL 537  
QY 508 LLSGIVQOQNLLRAIEAQOHLQLTVWGIK-LQARVLAVERYLKDQQLLGIWGCSGKLI 566  
Db 538 LLSGIVQOQNLLRAIEAQOHLQLTVWGIKLQARVLAVERYLKDQQLLGIWGCSGKLI 597  
QY 567 CTTAVPWNASWNSKSLDQIWNNTMWEEREIDNVTLSLYIESQOQKNEQELLEL 626  
Db 598 CTTAVPWNASWNSKSLDKIWNNTMWEEREIDNVTLSLYIESQOQKNEQELLEL 657  
QY 627 DKWASLWNNFDISKWLWYIK 646  
Db 658 DKWASLWNNFDTIKWLWYIK 677

RESULT 3

US-09-492-739-28



QY 1 MRVKGIRKNOYHQLWRGGTLLGLMLMICSAREKLVWTVVYGVVPMKEATTTTLFCASDAKAY 60  
DB 1 MRVKGIRKNOYHQLWRGGTLLGLMLMICSAREKLVWTVVYGVVPMKEATTTTLFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDPNPQBIIVLENTFENFMKNNMVEQMHEDIISLWDQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDPNPQBIIVLENTFENFMKNNMVEQMHEDIISLWDQSLKPCVK 120  
QY 121 LPLCVTLHCTNLC-----NATNTKSSNKWKMDRGEIKNCSFKVGAG----- 162  
DB 121 LPLCVTLHCTNLC-----NATNTKSSNKWKMDRGEIKNCSFKVGAG----- 162  
QY 163 -----KLJNCNTSVITQACPKVSPFPIPHYCAPAGFALLKNDKFKNGSPC 210  
DB 163 -----KLJNCNTSVITQACPKVSPFPIPHYCAPAGFALLKNDKFKNGSPC 210  
QY 211 TNSVSTVQCTHGIRPVVSTOLLNGSLAEEGVVIRSENFDTNKTIIIVQLKESVEINCTRP 270  
DB 211 TNSVSTVQCTHGIRPVVSTOLLNGSLAEEGVVIRSENFDTNKTIIIVQLKESVEINCTRP 270  
QY 240 TNSVSTVQCTHGIRPVVSTHLLNGSLAEEGVVIRSENFDTNKTIIIVQLKEAVEINCTRP 299  
DB 240 TNSVSTVQCTHGIRPVVSTHLLNGSLAEEGVVIRSENFDTNKTIIIVQLKEAVEINCTRP 299  
QY 271 NNNTKRSITIGGRAFYATGDIIGDROAHCHNISGEKNNNTLKQIVTKLQAOFGNKTIVF 330  
DB 271 NNNTKRSITIGGRAFYATGDIIGDROAHCHNISGEKNNNTLKQIVTKLQAOFGNKTIVF 330  
QY 300 NNNTKRSITIGGRAFYATGDIIGDROAHCHNISRAKNNNTLKQIVTKLQAOFGNKTIVF 359  
DB 300 NNNTKRSITIGGRAFYATGDIIGDROAHCHNISRAKNNNTLKQIVTKLQAOFGNKTIVF 359  
QY 331 KOSGSDPEIVMHSNCGGEFFVCNSTQLFNSNTWNTTIGPNNTG--TITLPCRIKQIIN 388  
DB 331 KOSGSDPEIVMHSNCGGEFFVCNSTQLFNSNTWNTTIGPNNTG--TITLPCRIKQIIN 388  
QY 360 NRSSGSDPEIVMHSNCGGEFFVCNSTQLFNSNTWNTTIGPNNTG--TITLPCRIKQIIN 419  
DB 360 NRSSGSDPEIVMHSNCGGEFFVCNSTQLFNSNTWNTTIGPNNTG--TITLPCRIKQIIN 419  
QY 389 RQEVGKAMYPPIRGQIRCSNITGLLITRDGQKEISNT-----TEIFRPGGDMRDNR 444  
DB 389 RQEVGKAMYPPIRGQIRCSNITGLLITRDGQKEISNT-----TEIFRPGGDMRDNR 444  
QY 420 MQEVGKAMYPPIRGQIRCSNITGLLITRDGQKEISNT-----TEIFRPGGDMRDNR 479  
DB 420 MQEVGKAMYPPIRGQIRCSNITGLLITRDGQKEISNT-----TEIFRPGGDMRDNR 479  
QY 445 SELYKVKVKBPLGVAFTKARVVQREKRAV--TLGAMFLGFLGAAGTGMARSJTLTV 503  
DB 445 SELYKVKVKBPLGVAFTKARVVQREKRAV--TLGAMFLGFLGAAGTGMARSJTLTV 503  
QY 480 SELYKVKVKBPLGVAFTKARVVQREKRAV--TLGAMFLGFLGAAGTGMARSJTLTV 539  
DB 480 SELYKVKVKBPLGVAFTKARVVQREKRAV--TLGAMFLGFLGAAGTGMARSJTLTV 539  
QY 504 QAROLLGIVQOONNLLRAEQAQHLQLTVWGK--LQARVLAVERYLKDQQLLGWGS 562  
DB 504 QAROLLGIVQOONNLLRAEQAQHLQLTVWGK--LQARVLAVERYLKDQQLLGWGS 562  
QY 540 QAROLLGIVQOONNLLRAEQAQHLQLTVWGK--LQARVLAVERYLKDQQLLGWGS 599  
DB 540 QAROLLGIVQOONNLLRAEQAQHLQLTVWGK--LQARVLAVERYLKDQQLLGWGS 599  
QY 563 GKLICTTAVPMNASNKSILQIWNMTWMEEREIDNTNLIYTLIBESQOQEKNEQE 622  
DB 563 GKLICTTAVPMNASNKSILQIWNMTWMEEREIDNTNLIYTLIBESQOQEKNEQE 622  
QY 600 GKLICTTAVPMNASNKSILQIWNMTWMEEREIDNTNLIYTLIBESQOQEKNEQE 659  
DB 600 GKLICTTAVPMNASNKSILQIWNMTWMEEREIDNTNLIYTLIBESQOQEKNEQE 659  
QY 623 LLELDKASLWNNFDSKWLWYIK 646  
DB 623 LLELDKASLWNNFDSKWLWYIK 646  
QY 660 LLELDKASLWNNFDSKWLWYIK 683  
DB 660 LLELDKASLWNNFDSKWLWYIK 683

## RESULT 5

US-08-022-835-6

; Sequence 6, Application US/08022835

; Patent No. 542030

; GENERAL INFORMATION:

; APPLICANT: Reitz Jr., Marvin S.

; APPLICANT: Franchini, Genoveffa

; APPLICANT: Markham, Phillip D.

; APPLICANT: Gallo, Robert C.

; APPLICANT: Leri, Franco C.

; APPLICANT: Popovic, Mikulas

; APPLICANT: Gartner, Suzanne

; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CUSHMAN, DARBY &amp; CUSHMAN

; STREET: Eleventh Floor, 1615 L. Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20036-5601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/022.835  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/599,491  
; FILING DATE: 17-OCT-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Scott, Watson T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 855 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-022-835-6

Query Match 85.7%; Score 2978.5; DB 1; Length 855;

Best Local Similarity 82.3%; Pred. No. 8.2e-238;

Matches 561; Conservative 36; Mismatches 48; Indels 37; Gaps 5;

QY 1 MRVKGIRKNOYHQLWRGGTLLGLMLMICSAREKLVWTVVYGVVPMKEATTTTLFCASDAKAY 60  
DB 1 MRVTEIRKSYQHWRWGMILGILMICSAREKLVWTVVYGVVPMKEATTTTLFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDPNPQBIIVLENTFENFMKNNMVEQMHEDIISLWDQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDPNPQBIIVLENTFENFMKNNMVEQMHEDIISLWDQSLKPCVK 120  
QY 121 LPLCVTLHCTNLC-----NATNTKSSNKWKMDR-----GEIKNCSFKVGAG----- 162  
DB 121 LPLCVTLHCTNLC-----NATNTKSSNKWKMDR-----GEIKNCSFKVGAG----- 162  
QY 163 -----KLJNCNTSVITQACPKVSPFPIPHYCAPAGFALLKNDKFKNGS 207  
DB 163 -----KLJNCNTSVITQACPKVSPFPIPHYCAPAGFALLKNDKFKNGS 207  
QY 181 KLDIAPIDNNNNRRLISCVITQACPKVSPFPIPHYCAPAGFALLKNDKFKNGK 240  
DB 181 KLDIAPIDNNNNRRLISCVITQACPKVSPFPIPHYCAPAGFALLKNDKFKNGK 240  
QY 208 GPCITVSTVQCTHGIRPVVSTOLLNGSLAEEGVVIRSENFDTNKTIIIVQLKESVEINC 267  
DB 208 GPCITVSTVQCTHGIRPVVSTOLLNGSLAEEGVVIRSENFDTNKTIIIVQLKESVEINC 267  
QY 241 GPCITVSTVQCTHGIRPVVSTOLLNGSLAEEGVVIRSENFDTNKTIIIVQLKESVEINC 300  
DB 241 GPCITVSTVQCTHGIRPVVSTOLLNGSLAEEGVVIRSENFDTNKTIIIVQLKESVEINC 300  
QY 268. TRPNNTKRSITIGGRAFYATGDIIGDROAHCHNISGEKNNNTLKQIVTKLQAOFGNKT 327  
DB 268. TRPNNTKRSITIGGRAFYATGDIIGDROAHCHNISGEKNNNTLKQIVTKLQAOFGNKT 327  
QY 301 TRPNNTKRSITIGGRAFYATGDIIGDROAHCHNISRAKNNNTLKQIVTKLQAOFGNKT 360  
DB 301 TRPNNTKRSITIGGRAFYATGDIIGDROAHCHNISRAKNNNTLKQIVTKLQAOFGNKT 360  
QY 328 IIVKOSGSDPEIVMHSNCGGEFFVCNSTQLFNSNTWNTTIGPNNT--NGTITLPCRIKQ 385  
DB 328 IIVKOSGSDPEIVMHSNCGGEFFVCNSTQLFNSNTWNTTIGPNNT--NGTITLPCRIKQ 385  
QY 361 IIVKOSGSDPEIVMHSNCGGEFFVCNSTQLFNSNTWNTTIGPNNT--NGTITLPCRIKQ 420  
DB 361 IIVKOSGSDPEIVMHSNCGGEFFVCNSTQLFNSNTWNTTIGPNNT--NGTITLPCRIKQ 420  
QY 386 IINRWQEVGKAMYPPIRGQIRCSNITGLLITRDGQKEISNTTEIFRPGGDMRDNR 445  
DB 386 IINRWQEVGKAMYPPIRGQIRCSNITGLLITRDGQKEISNTTEIFRPGGDMRDNR 445  
QY 421 IINRWQEVGKAMYPPIRGQIRCSNITGLLITRDGQKEISNTTEIFRPGGDMRDNR 479  
DB 421 IINRWQEVGKAMYPPIRGQIRCSNITGLLITRDGQKEISNTTEIFRPGGDMRDNR 479  
QY 446 ELYKYKVKVKBPLGVAFTKARVVQREKRAV--TLGAMFLGFLGAAGTGMARSJTLTV 505  
DB 446 ELYKYKVKVKBPLGVAFTKARVVQREKRAV--TLGAMFLGFLGAAGTGMARSJTLTV 505  
QY 480 ELYKYKVKVKBPLGVAFTKARVVQREKRAV--TLGAMFLGFLGAAGTGMARSJTLTV 539  
DB 480 ELYKYKVKVKBPLGVAFTKARVVQREKRAV--TLGAMFLGFLGAAGTGMARSJTLTV 539  
QY 506 ROLLGIVQOONNLLRAEQAQHLQLTVWGK--LQARVLAVERYLKDQQLLGWGS 564  
DB 506 ROLLGIVQOONNLLRAEQAQHLQLTVWGK--LQARVLAVERYLKDQQLLGWGS 564  
QY 540 ROLLGIVQOONNLLRAEQAQHLQLTVWGK--LQARVLAVERYLKDQQLLGWGS 599  
DB 540 ROLLGIVQOONNLLRAEQAQHLQLTVWGK--LQARVLAVERYLKDQQLLGWGS 599  
QY 565 LICTTAVPMNASNKSILQIWNMTWMEEREIDNTNLIYTLIBESQOQEKNEQE 624  
DB 565 LICTTAVPMNASNKSILQIWNMTWMEEREIDNTNLIYTLIBESQOQEKNEQE 624



Db 600 LICITAVPNASWSNLSKINWDMTWIEWDREINNTYSIIYSLIEESQOQEKNEQELL 659

QY 625 ELDKWSLWNPFDISKWLWYIK 646  
|||||:|||||

Db 660 ELDKWSLWNPFDITKWLWYIK 681

RESULT 6

US-08-388-809-6

; Sequence 6, Application US/08388809

; Patent No. 5576000

; GENERAL INFORMATION:

; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,

; APPLICANT: GENOVEFFA; MARKHAM, PHILIP D. GALLO, ROBERT

; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND

; APPLICANT: GARTNER, SUZANNE

; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK, 3.5"

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/388,809

; FILING DATE: 15-FEB-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/022,835

; FILING DATE: 25-FEB-1993

; APPLICATION NUMBER: US 07/599,491

; FILING DATE: 17-OCT-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: LESLIE A. SERUNIAN

; REGISTRATION NUMBER: 35,353

; REFERENCE/DOCKET NUMBER: 2026-4092US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 855 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-388-809-6

Query Match 85.7%; Score 2978.5; DB 1; Length 855;

Best Local Similarity 82.3%; Pred. No. 8.2e-238;

Matches 561; Conservative 36; Mismatches 48; Indels 37; Gaps 5;

QY 1 MRVKGIRKQYQHLWGGTLLGLMLICSAVEKLVTVTVYGVVPWKEATTTLCASDAKAY 60  
|||||:|||||

Db 1 MRVTEIRKSYQHWRRWGMGLLGLMLICNAEELKLVTVTVYGVVPWKEATTTLCASDRKAY 60  
|||||:|||||

QY 61 DTEVHNWATHACVPTDPNPQELVLENVTENFMNKNVQEHEDIISLWQSLKPCVK 120  
|||||:|||||

Db 61 DTEVHNWATHACVPTDPNPQELVLENVTENFMNKNVQEHEDIISLWQSLKPCVK 120  
|||||:|||||

QY 121 LTPLCVTLHCTLNKATNTKSNWKEMDR-----GEIKNCSFKVGAG----- 162  
|||||:|||||

Db 121 LTPLCVTLNCTDLRNATNGNDNTTSSSRGMVGGGEMKNCSFNITNIRGKVQKEALFY 180  
|||||:|||||

QY 163 -----KLINCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNCKKFNKS 207  
:|||||:|||||

Db 181 KLDIAPINNNNSNRVRLISCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNCKKFNK 240

QY 208 GECTNVSTVQCTHGIRPVVSTOLLNGSLAEGGVVIRSENFTDNAKTIIVQLKESVEINC 267  
|||||:|||||

Db 241 GECTNVSTVQCTHGIRPVVSTOLLNGSLAEGGVVIRSFANFADNAKVIIVQLNESVEINC 300  
|||||:|||||

QY 268 TRPNNTKRSITIGPGRAPYATGDIIGDIROAHCHNISGEKWNNTLKQIVTKLOAQFGNKT 327  
|||||:|||||

Db 301 TRPNNTKRSIHIGPGRAPYTTGTEIIGDIROAHCHNLRAKWNNTLNKIVIKLREQFGNKT 360  
|||||:|||||

QY 328 IVFKOSSGDDPEIVMHSFNCGGEFFVCNSTOLFNSTWNTTNGTITLPCRKQ 385  
|||||:|||||

Db 361 IVFKHSSGDDPEIVTHSFNCGGEFFVCNSTOLFNSTWNTTNGTITLPCRKQ 420  
|||||:|||||

QY 386 IINRWQEVGKAMYAPPPIRGQIRCCSNITGLLITRQGGKEISNTTIFRPGGDMRDNWS 445  
|||||:|||||

Db 421 IINRWQEVGRAMYAPPPIRGQIRCCSNITGLLITRQGGPE-DNKTEVFRPGGDMRDNWS 479  
|||||:|||||

QY 446 ELYKYKVKIPLGVAPTAKRRVVQREKRAVTLGAMFLGFLGAGSTNGARSLTLTVQA 505  
|||||:|||||

Db 480 ELYKYKVKIPLGVAPTAKRRVVQREKRAVIGAVFLGFLGAGSTNGAAAMTLTVQA 539  
|||||:|||||

QY 506 ROLLSGIVQQNNLRAIEAOOHLQLTVMGK-LQARVLAVERYLKOOHLGIWCCSGK 564  
|||||:|||||

Db 540 ROLLSGIVQQNNLRAIEAOOHLQLTVMGK-LQARVLAVERYLRDQQLGIWCCSGK 599  
|||||:|||||

QY 565 LICITAVPNASWSNLSKIDQIWNNTWMEWEIEDNYTNLIYTLIEESQOQEKNEQELL 624  
|||||:|||||

Db 600 LICITAVPNASWSNLSKLNKIDWNNTWIEWDREINNTYSIIYSLIEESQOQEKNEQELL 659  
|||||:|||||

QY 625 ELDKWSLWNPFDISKWLWYIK 646  
|||||:|||||

Db 660 ELDKWSLWNPFDITKWLWYIK 681  
|||||:|||||

RESULT 7

US-08-647-714-6

; Sequence 6, Application US/08647714

; Patent No. 5869313

; GENERAL INFORMATION:

; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,

; APPLICANT: GENOVEFFA; MARKHAM, PHILIP D. GALLO, ROBERT

; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND

; APPLICANT: GARTNER, SUZANNE

; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK, 3.5"

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/647,714

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/388,809

; FILING DATE: 15-FEB-1995

; APPLICATION NUMBER: US 08/022,835

; FILING DATE: 25-FEB-1993

; APPLICATION NUMBER: US 07/599,491

; FILING DATE: 17-OCT-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: LESLIE A. SERUNIAN

; REGISTRATION NUMBER: 35,353

; REFERENCE/DOCKET NUMBER: 2026-4092US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 855 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-647-714-6

Query Match 85.7%; Score 2978.5; DB 1; Length 855;

Best Local Similarity 82.3%; Pred. No. 8.2e-238;

Matches 561; Conservative 36; Mismatches 48; Indels 37; Gaps 5;

QY 1 MRVKGIRKQYQHLWGGTLLGLMLICSAVEKLVTVTVYGVVPWKEATTTLCASDAKAY 60  
|||||:|||||

Db 1 MRVTEIRKSYQHWRRWGMGLLGLMLICNAEELKLVTVTVYGVVPWKEATTTLCASDRKAY 60  
|||||:|||||

QY 61 DTEVHNWATHACVPTDPNPQELVLENVTENFMNKNVQEHEDIISLWQSLKPCVK 120  
|||||:|||||

Db 61 DTEVHNWATHACVPTDPNPQELVLENVTENFMNKNVQEHEDIISLWQSLKPCVK 120  
|||||:|||||

QY 121 LTPLCVTLHCTLNKATNTKSNWKEMDR-----GEIKNCSFKVGAG----- 162  
|||||:|||||

Db 121 LTPLCVTLNCTDLRNATNGNDNTTSSSRGMVGGGEMKNCSFNITNIRGKVQKEALFY 180  
|||||:|||||

QY 163 -----KLINCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNCKKFNKS 207  
:|||||:|||||

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-647-714-6

Query Match      85.7%; Score 2978.5; DB 2; Length 855;
Best Local Similarity 82.3%; Pred. No. 8.2e-238;
Matches 561; Conservative 36; Mismatches 48; Indels 37; Gaps 5;

QY 1 MRVGIKKNYQHLWRGGTLLGLMLTCSAVEKLWTVVYGVVPWKAEATTLFCASDAKAY 60
DB 1 MRVTEIRKSYQHWRMGIMLLGLMTCNABEKLWTVVYGVVPWKAEATTLFCASDRKAY 60
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENTNFNMKNMKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEVHNWATHACVPTDPNPQEVLEKLVNTNFNMKNMKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPCLVTLHCTNLKNATNTKSNWKMEDR-----GEIKNCSPKVGAG----- 162
DB 121 LTPCLVTLNCTDLRNATNGNDNTTSSRGMVGGMKNCSTNIITIRGKVQKEVALPY 180
QY 163 -----KLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILLKCNDDKFKNGS 207
DB 181 KLDIAPIDNNSNRRLISNTSVITQACPKVSFEPIPIHYCAPAGFAILLKCNDDKFKNGK 240
QY 208 GCTNVSTVQCTHGIRPVVSTQLLNGSLAEBGVIRSENFNTDNAKTIIVQLKESVEINC 267
DB 241 GCTNVSTVQCTHGIRPVVSTQLLNGSLAEBEWIRSFANFADNAKTIIVQLNESVEINC 300
QY 268 TRPNNTRKSIITGPGRAFYATGDIIGDIRQAHNCISGEKNNTLKQIVTKLQAOFGNKT 327
DB 301 TRPNNTRKSIHGPGRAPYTTGDIIGDIRQAHNCISRAKNDTLNKIVIKLREQFGNKT 360
QY 328 IVFKOSSGGDPEIVMHSFNCGGEFFYCNSQLFNSNTWNNTIGPNNNT--NGTITLPCRIRK 385
DB 361 IVFKSHSSGGDPEIVTHSPFCGGEFFYCNSQLFNSNTWNNTVENNTITLPCRIRK 420
QY 386 IINRWQEVGKAMYAPPIRGQIRCSSNITGLLTRDGGKEISNTTEIFRPGGDMRDNRWS 445
DB 421 IINRWQEVGRAMYAPPIRGQIRCSSNITGLLTRDGGPE--DNKTEVFRPGGDMRDNRWS 479
QY 446 ELYKVKVVKIEPLGVAPTAKRRVQREKRAVTLGAMFLGLGAAGSTMGARSLTLTVQA 505
DB 480 ELYKVKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGLGAAGSTMGAAWTLTVQA 539
QY 506 ROLLSGIVQOQNLLRAIEAQHLLQLTWGIK--LOARVLAVERYLKQOLLGIWGCSGK 564
DB 540 RLLLSGIVQOQNLLRAIEAQHLLQLTWGIKQLQARVLAVERYLRDQOLLGIWGCSGK 599
QY 565 LICTTAPVNASNSKSLDOIKNMTWMEEREIDNYNLIVLTIESONQOEKQOEILL 624
DB 600 LICTTAPVNASNSKSLNKIDWNTWIEWDRINNTYIIISLTIESONQOEKQOEILL 659
QY 625 ELDKWSLWNPFDISKWLWYIK 646
DB 660 ELDKWSLWNPFDITKWLWYIK 681

```

```

RESULT 8
US-07-956-483-15
; Sequence 15, Application US/07956483
; Patent No. 6261799
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; TITLE OF INVENTION: gp160 VARIANT

```

```

; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,483
; FILING DATE: 31-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/19742
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 05392
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-956-483-15

```

```

Query Match      85.7%; Score 2976; DB 3; Length 855;
Best Local Similarity 82.3%; Pred. No. 1.3e-237;
Matches 562; Conservative 34; Mismatches 49; Indels 38; Gaps 6;

QY 1 MRVGIKKNYQHLWRGGTLLGLMLTCSAVEKLWTVVYGVVPWKAEATTLFCASDAKAY 60
DB 1 MKVGTTRRYQHLWRGWTLLGLMLTCSATEKLWTVVYGVVPWKAEATTLFCASDARAY 60
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENTNFNMKNMKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEVHNWATHACVPTDPNPQEVLEKLVNTNFNMKNMKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPCLVTLHCTNLKNATNTKSNWKMEDRGEIKNCSEFNITTSIRDKIQENALFRNLDVV 180
DB 121 LTPCLVTLNCTDLGKATNTNSNWKEEIKGEIKNCSEFNITTSIRDKIQENALFRNLDVV 180
QY 163 -----KLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILLKCNDDKFKNGS 208
DB 181 PIDNASTTNTYNYRLIHCNRSVITQACPKVSFEPIPIHYCTPAGFAILLKCNKTKFNGK 240
QY 209 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEBGVIRSENFNTDNAKTIIVQLKESVEINC 268
DB 241 TCTNVSTVQCTHGIRPVVSTQLLNGSLAEBEVRSDNFNTNNAKTIIVQLNESVAINCT 300
QY 269 RPNNTKRSITIGPGRAFYATGDIIGDIRQAHNCISGEKNNTLKQIVTKLQAOFG--NKT 327
DB 301 RPNNTKRSIYGPGRAFHTTGRIIGDIRKAHCNISRQMNNTLEQIVKLRQFGNNTK 360
QY 328 IVFKOSSGGDPEIVMHSFNCGGEFFYCNSQLFNSNTWNNTIGPNNNT--NNTIGPNNNTGTTLPCRIRK 385
DB 361 IVFNQSSGGDPEIVMHSFNCRCGEFFYCNTQLPNTWRLNHTEGTKG--NDTITLPCRIRK 419
QY 386 IINRWQEVGKAMYAPPIRGQIRCSSNITGLLTRDGGKEISNTTEIFRPGGDMRDNRWS 445

```

Db 420 IINMWOEVCKAMYAPPIGQISCSNITGLLLTRDGGTNVTDTEVFRPGGDMRDNRWS 479

QY 446 ELYKYKVKIEPLGVAPTAKRRVVQREKRAV-TLGAMFLGFLGAAGSTMGARSITLTVQ 504

Db 480 ELYKYKVKIEPLGIAPTAKRRVVQREKRAVGIIVGAMFLGFLGAAGSTMGAVSLTLTVQ 539

QY 505 ARQLLSGIVQOQNLLRAIEAOQHLLQLTWGIK-LOARVLAVERYLKDQQLLGIWGCSSG 563

Db 540 ARQLLSGIVQOQNLLRAIEAOQHLLQLTWGIKQLOARVLAVERYLRDQQLLGIWGCSSG 599

QY 564 KLICTTAVPWNASWSNKSLDQIWNMTWMEWEREDINYTNLYTLIEESQNOQEKNEOBL 623

Db 600 KLICTTAVPWNASWSNKSLEDIWDNMTWQWEREDINYTNLYTLIEESQNOQEKNEOBL 659

QY 624 LELDKWASLWNNWFDISKWLWYIK 646

Db 660 LELDKWASLWNNWFSITNWLWYIK 682

RESULT 9

US-08-448-603A-30

; Sequence 30, Application US/08448603A

; Patent No. 5864027

; GENERAL INFORMATION:

; APPLICANT: Berman, Phillip W.

; APPLICANT: Nakamura, Gerald R.

; TITLE OF INVENTION: HIV Envelope Polypeptides

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP

; STREET: 3 Embarcadero Center

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/448,603A

; FILING DATE: 07-JUN-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/072,833

; FILING DATE: 07-JUN-93

; ATTORNEY/AGENT INFORMATION:

; NAME: Haliday, Emily

; REGISTRATION NUMBER: 38903

; REFERENCE/DOCKET NUMBER: 14918-704

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-393-2000

; TELEFAX: 415-393-2286

; TELEX:

; INFORMATION FOR SEQ ID NO: 30:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 857 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-448-603A-30

Query Match 85.6%; Score 2974; DB 237; Length 857;

Best Local Similarity 81.2%; Pred. No. 1.9e-237;

Matches 556; Conservative 43; Mismatches 46; Indels 40; Gaps 4;

QY 1 MRVKGIRKNYQHLWRGGTLLGLMLICSAVEKLWTVVYGVVPWKAEATTLTFCASDAKAY 60

Db 1 MRVKGIRKNYQHLWRGGTLLGLMLICSAAGKLWTVVYGVVPWKETTLTFCASDAKAY 60

QY 61 DTEHNVWATHACVPTDNPQIBVLNTNFNMWKNMNVQEHEDIIISLWDQSLKPCVK 120

Db 61 DTEHNVWATHACVPTDNPQIBVLNTNFNMWKNMNVQEHEDIIISLWDQSLKPCVK 120

QY 121 LTPCLVTLHCNTNLKNATNTKSSNMKEMDRGEIKNGSKFKVGAG----- 162

Db 121 LTPCLVTLNCDDAGNTTNTSSSRKLEKGEIKNGSFNITTSVRDKMQKETALFNKLDIV 180

QY 163 -----KLINCNTSVITQACPKVSEPIPIHYCAPAGPAIFAILKCNCKKFNK 206

Db 181 PIDDDNRNSTRNSTNYRLISCNSTVITQACPKVSEPIPIHFCTPAGFALLCKNKTENG 240

QY 207 SGPCTNVSTVOCTHGIRPVVSTQLLNGSLAEGVVRSENFETDNAKTIIVOLKESVEIN 266

Db 241 SGPCKNVSTVOCTHGIRPVVSTQLLNGSLAEGVVRSENFETNNAKTIIVQLTBPVKIN 300

QY 267 CTRPNNTNRKSIITIGPGRFAYATGDIIGDIRAHNCISGEKWNNTLKQIVTKLOAOFGNK 326

Db 301 CTRPNNTNRKSIPIGPGRFAYATGDIIGNIROAHCNLSRTDWNNTLGQIVKLRQFGNK 360

QY 327 TIVFKQSSGGDPEIIVMHSFNCGBPFYCNSTQLFNSTWNNTIGPNNT-----NGTITLPCR 382

Db 361 TIIFNHSSGGDPEIIVMHSFNCGBPFYCNSTQLFSDTWDNTKVSNGTSTEENSTITLPCR 420

QY 383 IKQIINRWQEVCKAMYAPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGDMRDN 442

Db 421 IKQIVNMWQEVCKAMYAPPIRGQIRCSNITGLLLTRDGGSN-NSMNETFRPGGDMRDN 479

QY 443 WRSELYKYKVKIEPLGVAPTAKRRVVQREKRAVTLGAMFLGFLGAAGSTMGARSITLT 502

Db 480 WRSELYKYKVKIEPLGVAPTAKRRVVQREKRAVIGANVFLGFLGAAGSTMGARSITLT 539

QY 503 VOARQLLSGIVQOQNLLRAIEAOQHLLQLTWGIK-LOARVLAVERYLKDQQLLGIWGC 561

Db 540 VOARQLLSGIVQOQNLLRAIEAOQHLLQLTWGIKQLOARVLAVERYLRDQQLLGIWGC 599

QY 562 SGKLICTTAVPWNASWSNKSLDQIWNMTWMEWEREDINYTNLYTLIEESQNOQEKNEO 621

Db 600 SGKLICTTAVPWNASWSNKSLEDIWDNMTWMEWEREDINYTNLYTLIEESQNOQEKNEO 659

QY 622 ELLELDKWASLWNNWFDISKWLWYIK 646

Db 660 DLELDQWASLWNNWFSITNWLWYIK 684

RESULT 10

US-09-134-075-30

; Sequence 30, Application US/09134075

; Patent No. 6042836

; GENERAL INFORMATION:

; APPLICANT: Berman, Phillip W.

; APPLICANT: Nakamura, Gerald R.

; TITLE OF INVENTION: HIV Envelope Polypeptides

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP

; STREET: 3 Embarcadero Center

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/134,075

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/448,603

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Haliday, Emily

```

; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX:
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-134-075-30

Query Match      85.6%; Score 2974; DB 3; Length 857;
Best Local Similarity 81.2%; Pred. No. 1.9e-237;
Matches 556; Conservative 43; Mismatches 46; Indels 40; Gaps 4;

QY 1 MRVKGIRKRYQHWRGGTLLGLMLICSAVEKLWTVVYGVVPMWKEATTTLCASDAKAY 60
DB 1 MRVKGIRKRYQHWRGGTLLGLMLICSAAGKLVTVVYGVVPMWKEATTTTLFCASDAKAY 60
QY 61 DTEVNVWATHACVPTDNPQBIIVLENTENFMNKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEVNVWATHACVPTDNPQBIIVLENTENFMNKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPCLVTLLCTNLKNATNTKSSNWKEMDRGEIKNGCFKVGAG----- 162
DB 121 LTPCLVTLLCTNLKNATNTKSSNWKEMDRGEIKNGCFNITTSVRDKMQKETALFNKLDIV 180
QY 163 -----KLINCNTSVITQACPKVSEPIPIHYCAPAGFAILKCNCKKNG 206
DB 163 PIDDNDNRSTRNSTRYRLISCNSTVITQACPKVSEPIPIHCTPAGFALLKCNCKKNG 240
QY 207 SGPCNTVSVQCTHGRPVVSTQLLNGSLABEGVIRSENFTDNKATIIIVOLKESVEIN 266
DB 241 SGPCKNVSVQCTHGRPVVSTQLLNGSLABEGVIRSENFTNNAKTIIVQLTEPVKIN 300
QY 267 CTRPNNTKRSITIGPGRAFYATGDIIGDIAHNCISGEKWNNTLKQIVTKLQAOFGNK 326
DB 301 CTRPNNTKRSIPIGPGRAFYATGDIIGNIAHNCNLSTDNNTLQGVKEKLRQFGNK 360
QY 327 TIVFKQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNNTIGPNNNT-----NGTITLPCR 382
DB 361 TIIFNHSSGGDPEIVMHSFNCGERFFYCNSTQLFSDTNDTKVSNGTSTEENSTITLPCR 420
QY 383 IKQIINRWQEVGKAMYAPPPIRGIRCSNITGLLITRDGCKEISNTTEIFRGGGMDRN 442
DB 421 IKQIVNMWQEVGKAMYAPPPIRGIRCSNITGLLITRDGGSN-NSMNETFRGGGMDRN 479
QY 443 WSELYKVKVTEPLGVAPTAKRRVQREKRAVTLGAMFLGLGAAGSTWGAASLTLT 502
DB 480 WSELYKVKVTEPLGVAPTAKRRVQREKRAVTLGAVLGLGAAGSTWGAASLTLT 539
QY 503 VQARQLLSGIVQQNNLLRAIEAQHLLQLTVWGIK-LQARVLAVERYLKDQQLGIWGC 561
DB 540 VQARQLLSGIVQQNNLLRAIEAQHLLQLTVWGIKLQARVLAVERYLRDQQLGIWGC 599
QY 562 SGKLICTTAVPNWASNSKSLDQIWNMTWMEBERIDNYTLIYTLIBESONQOEKNEQ 621
DB 600 SGKLICTTAVPNWASNSKSLDQIWNMTWMEBERIENYTLIYTLIBESONQOEKNEQ 659
QY 622 ELLELDKWSLWNPFDISKWLAVIK 646
DB 660 DLELDQWASLWNPFDISKWLAVIK 684

```

RESULT 11

US-09-492-739-30  
; Sequence 30, Application US/09492739  
; Patent No. 6331404  
; GENERAL INFORMATION:  
; APPLICANT: Berman, Phillip W.

```

; Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/492,739
; FILING DATE: 27-Jan-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/134,075
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haliday, Emily
; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX: <Unknown>
;
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-492-739-30

Query Match      85.6%; Score 2974; DB 4; Length 857;
Best Local Similarity 81.2%; Pred. No. 1.9e-237;
Matches 556; Conservative 43; Mismatches 46; Indels 40; Gaps 4;

QY 1 MRVKGIRKRYQHWRGGTLLGLMLICSAVEKLWTVVYGVVPMWKEATTTLCASDAKAY 60
DB 1 MRVKGIRKRYQHWRGGTLLGLMLICSAAGKLVTVVYGVVPMWKEATTTTLFCASDAKAY 60
QY 61 DTEVNVWATHACVPTDNPQBIIVLENTENFMNKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEVNVWATHACVPTDNPQBIIVLENTENFMNKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPCLVTLLCTNLKNATNTKSSNWKEMDRGEIKNGCFKVGAG----- 162
DB 121 LTPCLVTLLCTNLKNATNTKSSNWKEMDRGEIKNGCFNITTSVRDKMQKETALFNKLDIV 180
QY 163 -----KLINCNTSVITQACPKVSEPIPIHYCAPAGFAILKCNCKKNG 206
DB 163 PIDDNDNRSTRNSTRYRLISCNSTVITQACPKVSEPIPIHCTPAGFALLKCNCKKNG 240
QY 207 SGPCNTVSVQCTHGRPVVSTQLLNGSLABEGVIRSENFTDNKATIIIVOLKESVEIN 266
DB 241 SGPCKNVSVQCTHGRPVVSTQLLNGSLABEGVIRSENFTNNAKTIIVQLTEPVKIN 300
QY 267 CTRPNNTKRSITIGPGRAFYATGDIIGDIAHNCISGEKWNNTLKQIVTKLQAOFGNK 326
DB 301 CTRPNNTKRSIPIGPGRAFYATGDIIGNIAHNCNLSTDNNTLQGVKEKLRQFGNK 360
QY 327 TIVFKQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNNTIGPNNNT-----NGTITLPCR 382
DB 361 TIIFNHSSGGDPEIVMHSFNCGERFFYCNSTQLFSDTNDTKVSNGTSTEENSTITLPCR 420
QY 383 IKQIINRWQEVGKAMYAPPPIRGIRCSNITGLLITRDGCKEISNTTEIFRGGGMDRN 442

```

Db 421 IKQIVNMWQEVGKAMYPPIRGQIRCSSNITGLLLTRDGGSN-NSMNETTFRGGGDMRDN 479  
Qy 443 WSELYKYKVKIEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAAGSTMGARSILTLT 502  
Db 480 WSELYKYKVKIEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAAGSTMGARSILTLT 539  
Qy 503 VOAROLLSGIVQOONLLRAIAQOHLLOLTWGIK-LOARVLAVERYLKDQOLLGIWGC 561  
Db 540 VOAROLLSGIVQOONLLRAIAQOHLLOLTWGIK-LOARVLAVERYLKDQOLLGIWGC 599  
Qy 562 SKGLICTTAVPNWASNSKSLDQIWNMTWMEWEREIDNYTLIYTLIEESNQOEKNEQ 621  
Db 600 SKGLICTTAVPNWASNSKSLDQIWNMTWMEWEREIDNYTLIYTLIEESNQOEKNEQ 659  
Qy 622 ELLELDKWSLWNFWDISKWLYIK 646  
Db 660 DLELDQWASLWNFWSITKWLWYIK 684

RESULT 12

US-07-956-483-13  
; Sequence 13, Application US/07956483  
; Patent No. 6261799  
; GENERAL INFORMATION:  
; APPLICANT: KIENY, Marie-Paule  
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
; TITLE OF INVENTION: gp160 VARIANT  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/956,483  
; FILING DATE: 31-DEC-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 92/19742  
; FILING DATE: 12-NOV-1992  
; APPLICATION NUMBER: FR 91 05392  
; FILING DATE: 02-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crane-Feury, Sharon E  
; REGISTRATION NUMBER: 36,113  
; REFERENCE/DOCKET NUMBER: 017753-005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 865 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-07-956-483-13

Query Match 84.8%; Score 2947; DB 3; Length 865;  
Best Local Similarity 80.1%; Pred. No. 3.4e-235;  
Matches 554; Conservative 40; Mismatches 52; Indels 46; Gaps 5;  
Qy 1 MRVKGIRKYNQHLWGGTLLGLMLICSAVEKLWTVYGVVWKEATTLFCASDAKAY 60  
Db 1 MRVMEMRKNCQHLWKGWTLGLMLICSAEDLWTVYGVVWKEATTLFCASEAKAY 60

Qy 61 DTEVHNWATHACVPTDNPQEIIVLENTENFNMKNMKNMVEQMHEDIISLWDSLKPCVK 120  
Db 61 KTEVHNWAKHACVPTDNPQEVLENTENFNMKNMKNMVEQMHEDIISLWDSLKPCVK 120  
Qy 121 LTPLCVTLHCT--NLKNATNTKSSNWKMDRGEIKNCSEFKV----- 159  
Db 121 LTPLCVTLNCTDANLNGNTVTSSSGGTWMGEIKNCSEFKV----- 180  
Qy 160 -----GAGKLINCNTSVITQACPQKVSFEPIPIHYCAPAGFAILK 198  
Db 181 VPIEKGNI SPKNTSNNTSYNYTLIHNSNVITQACPQKVSFEPIPIHYCTPAGFAILK 240  
Qy 199 CNDKFKNGSGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEAGVIRSENFTDNTAKTIIVQ 258  
Db 241 CNDKFKNGTGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEAGVIRSENFTDNTAKTIIVQ 300  
Qy 259 LKESVEINCTRPNNNTRKSIITIGPGRAPYATGDIIGDIRQAHNCISGEKWNNTLQIVTK 318  
Db 301 LNASVQINCTRPNNNTRKSIITIGPGRVYATGQIIGDIRKAHCNLSRAQMNNTLQVVTK 360  
Qy 319 LQAQFGNKTIVFKQSSGGDPEIVMHSFNCGGEFFYCNSTQLPFNSNTWNTTIGPNTNG--T 376  
Db 361 LREQFDNKTIVFTSSGGDPEIVLHSPNCGGFFYCNSTQLPFNSNTWNTTIGPNTNG--T 420  
Qy 377 ITLPCRIKQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRDGGKISINTTEIFRPGG 436  
Db 421 ITLPCRIKQIVNMWQEVGKAMYPPIRGQIRCSSNITGLLLTRDGGEDTTNTTEIFRLOG 480  
Qy 437 GDMRDWSELYKYKVKIEPLGVAPTAKRRVQREKRAV-TLGAMFLGFLGAAGSTMG 495  
Db 481 GDMRDWSELYKYKVKIEPLGVAPTAKRRVQREKRAV-TLGAMFLGFLGAAGSTMG 540  
Qy 496 ARSLTLTVQAROLLSGIVQOONLLRAIAQOHLLOLTWGIK-LOARVLAVERYLKDQO 554  
Db 541 AGSITLTVQARHLLSGIVQOONLLRAIAQOHLLOLTWGIK-LOARVLAVERYLKDQO 600  
Qy 555 LLGIWCSGKLCITCTTAVPNWASNSKSLDQIWNMTWMEWEREIDNYTLIYTLIEESQN 614  
Db 601 LLGIWCSGKLCITCTTAVPNWASNSKSLDQIWNMTWMEWEREIDNYTLIYTLIEESQN 660  
Qy 615 QOEKNEQELLEDKWSLWNFWDISKWLYIK 646  
Db 661 QOEKNEQELLEDKWSLWNFWDISKWLYIK 692

RESULT 13

US-09-124-900-9  
; Sequence 9, Application US/09124900  
; Patent No. 6268484  
; GENERAL INFORMATION:  
; APPLICANT: KATINGER, Hermann  
; APPLICANT: BUCHACHER, Andrea  
; APPLICANT: ERNST, Wolfgang  
; APPLICANT: BALLAUN, Claudia  
; APPLICANT: PURTSCHER, Martin  
; APPLICANT: TRKOLA, Alexandra  
; APPLICANT: PREDL, Renate  
; APPLICANT: SCHWATZ, Christine  
; APPLICANT: KLIMA, Annelies  
; APPLICANT: STEINDL, Franz  
; APPLICANT: MUSTER, Thomas  
; TITLE OF INVENTION: HIV-Vaccines  
; FILE REFERENCE: 1939-112p  
; CURRENT APPLICATION NUMBER: US/09/124,900  
; CURRENT FILING DATE: 1998-07-30  
; PRIOR APPLICATION NUMBER: PCT/EP95/01481  
; PRIOR FILING DATE: 1995-04-19  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 9  
; LENGTH: 856  
; TYPE: PRT

```

; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (79)..(184)
; NAME/KEY: PEPTIDE
; LOCATION: (326)..(400)
US-09-124-900-9

Query Match      84.6%; Score 2939.5; DB 3; Length 856;
Best Local Similarity 82.6%; Pred. No. 1.4e-234;
Matches 568; Conservative 34; Mismatches 39; Indels 47; Gaps 12;

QY 1 MRVKGIRKXNYOHL-----WRGGTLLGLMLMICSAREKLVVTVYGVVPVWKEATTLFCASD 56
Db 1 MRVK---EKYQHLWRGWRGWTMLGLMLMICSAREKLVVTVYGVVPVWKEATTLFCASD 57

QY 57 AKAYDEVHNVWATHACVPTDPNPQEI VLENTFNMMKNMKNMVEQMHEDIISLWQSLK 116
Db 58 AKAYDEVHNVWATHACVPTDPNPQEVVLNVTFNMMKNMKNMVEQMHEDIISLWQSLK 117

QY 117 PCVKLTPLCVTLHCTNLKNATKSNWKE-MDRGEIKNCSPKVG---GK-----163
Db 118 PCVKLTPLCVSLKCTDLKNDNTNSSSGRMIMEKEIKNCSPNISTIRGKQKEYAFY 177

QY 164 -----LINCNTSVITQACPKVSEPIPIHYCAPAGFAILKCNKKFNKSG 208
Db 178 KLDIIPIDNDTTSYTLTSCNTSVITQACPKVSEPIPIHYCAPAGFAILKCNKTFNGTG 237

QY 209 PCTNVSTVQCTGIRPVVSTQLLNGSLAEGBVIRSENFNTDNAKTIIIVQLKESVEINCT 268
Db 238 PCTNVSTVQCTGIRPVVSTQLLNGSLAEGBVIRSENFNTDNAKTIIIVQLKESVEINCT 297

QY 269 RNNNTRKSIITI--GPGAFYATGDIIGDIOAHNCISGEKWNNTLKQIVTKLQAFG-N 325
Db 298 RNNNTRKSIIRIQRGPGAFVITGK-IGNMROAHNCISRAKWNNTLKQIDSKLREQFGNN 356

QY 326 KTIIVFKQSSGGDPEIVMHSFNCGGFFYCNSTOLFNTWNN---TIGPNTNG--TITL 379
Db 357 KTIIVFKQSSGGDPEIVMHSFNCGGFFYCNSTOLFNTWNN---TIGPNTNG--TITL 416

QY 380 PCRIKQIINRWQEVGKAMYPPIRGICSSNITGLLLTRDGGKISNTTIFRPGGDM 439
Db 417 PCRIKQIINRWQEVGKAMYPPIRGICSSNITGLLLTRDGGNS--NNESEIFRPGGDM 475

QY 440 RDNWSELKYKVKIEPLGVPATKAKRVVQREKRAVTLGAMFLGFLGAAGSTWARS 499
Db 476 RDNWSELKYKVKIEPLGVPATKAKRVVQREKRAVTLGAMFLGFLGAAGSTWARS 535

QY 500 TLTVQARQLLSGIVQOQNLLRAIEAQHLLQTLTVMGIK-LQARVLAVERYLKDOQLGI 558
Db 536 TLTVQARQLLSGIVQOQNLLRAIEAQHLLQTLTVMGIKLQARVLAVERYLKDOQLGI 595

QY 559 WCGSGKLICTTAVPMNASNSKSLDQIMWNMTWMEWERIDNYTNLIYTLIESQNQOEK 618
Db 596 WCGSGKLICTTAVPMNASNSKSLDQIMWNMTWMEWERIDNYTNLIYTLIESQNQOEK 655

QY 619 NEQELLELDKWSLWVNFIDSKWLVIK 646
Db 656 NEQELLELDKWSLWVNFIDSKWLVIK 683

```

RESULT 14

```

US-08-463-210-11
; Sequence 11, Application US/08463210
; Patent No. 6001977
; GENERAL INFORMATION:
; APPLICANT: CHANG, Nancy T.
; APPLICANT: GALLO, Robert C.
; APPLICANT: WONG-STALL, Floreie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.

```

```

; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,210
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/693,866
; FILING DATE: 23-JAN-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/659,339
; FILING DATE: 10-OCT-1984
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4193US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 863 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: HTLV-III
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..863
; OTHER INFORMATION: /note= "env protein of HTLV-III"
US-08-463-210-11

```

```

Query Match      84.6%; Score 2939.5; DB 3; Length 863;
Best Local Similarity 82.6%; Pred. No. 1.4e-234;
Matches 568; Conservative 34; Mismatches 39; Indels 47; Gaps 12;

QY 1 MRVKGIRKXNYOHL-----WRGGTLLGLMLMICSAREKLVVTVYGVVPVWKEATTLFCASD 56
Db 8 MRVK---EKYQHLWRGWRGWTMLGLMLMICSAREKLVVTVYGVVPVWKEATTLFCASD 64

QY 57 AKAYDEVHNVWATHACVPTDPNPQEI VLENTFNMMKNMKNMVEQMHEDIISLWQSLK 116
Db 65 AKAYDEVHNVWATHACVPTDPNPQEVVLNVTFNMMKNMKNMVEQMHEDIISLWQSLK 124

QY 117 PCVKLTPLCVTLHCTNLKNATKSNWKE-MDRGEIKNCSPKVG---GK-----163
Db 125 PCVKLTPLCVSLKCTDLKNDNTNSSSGRMIMEKEIKNCSPNISTIRGKQKEYAFY 184

QY 164 -----LINCNTSVITQACPKVSEPIPIHYCAPAGFAILKCNKKFNKSG 208
Db 185 KLDIIPIDNDTTSYTLTSCNTSVITQACPKVSEPIPIHYCAPAGFAILKCNKTFNGTG 244

QY 209 PCTNVSTVQCTGIRPVVSTQLLNGSLAEGBVIRSENFNTDNAKTIIIVQLKESVEINCT 268
Db 245 PCTNVSTVQCTGIRPVVSTQLLNGSLAEGBVIRSENFNTDNAKTIIIVQLKESVEINCT 304

QY 269 RNNNTRKSIITI--GPGAFYATGDIIGDIOAHNCISGEKWNNTLKQIVTKLQAFG-N 325
Db 305 RNNNTRKSIIRIQRGPGAFVITGK-IGNMROAHNCISRAKWNNTLKQIDSKLREQFGNN 363

QY 326 KTIIVFKQSSGGDPEIVMHSFNCGGFFYCNSTOLFNTWNN---TIGPNTNG--TITL 379

```

Db 364 KTIIFKQSGGDEIIVTHSFNCGGEFFYCNSTQLFNSTWFSNWTSGNSTEGSDTITL 423  
Qy 380 PCRIKQIINRWQEVGKAMVAPPPIRGQIRCSNITGLLTRDGGKEISNTTEIFRPGGDM 439  
Db 424 PCRIKQIINRWQEVGKAMVAPPISQIRCSNITGLLTRDGGNS--NNESEIFRPGGDM 482  
Qy 440 RDNWRSELYKYKVKVKEPIGLVAPTAKRRVVOREKRAVTLGAMFLGFLGAAGSTMGARSL 499  
Db 483 RDNWRSELYKYKVKVKEPIGLVAPTAKRRVVOREKRAVTLGAMFLGFLGAAGSTMGAASM 542  
Qy 500 TLTVOARQLLSGIVQOONLLRAIEAQHLLQLTVMGIK--LOARVLAVERYLKDQQLLGI 558  
Db 543 TLTVOARQLLSGIVQOONLLRAIEAQHLLQLTVMGIK--LOARVLAVERYLKDQQLLGI 602  
Qy 559 WCGSGKLICTTAVPWNASNSKSLDQIWNMTWMEWERIDNYTNLIYTLIESQNQOEK 618  
Db 603 WCGSGKLICTTAVPWNASNSKSLDQIWNMTWMEWERIDNYTNLIYTLIESQNQOEK 662  
Qy 619 NEQELLELDKWSLWNFIDISKLWYIK 646  
Db 663 NEQELLELDKWSLWNFITNLWYIK 690

RESULT 15

US-08-463-028-11  
; Sequence 11, Application US/08463028  
; Patent No. 6610476  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, Nancy T.  
; APPLICANT: GALLO, Robert C.  
; APPLICANT: WONG-STAAU, Flossie  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,028  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/693,866  
; FILING DATE: 23-JAN-1985  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/659,339  
; FILING DATE: 10-OCT-1984  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Serunian, Leslie A.  
; REGISTRATION NUMBER: 35,353  
; REFERENCE/DOCKET NUMBER: 2026-4193US3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 863 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: HTLV-III  
; FEATURE:

; NAME/KEY: Protein  
; LOCATION: 1..863  
; OTHER INFORMATION: /note= "env protein of HTLV-III"  
US-08-463-028-11  
Query Match 84.6%; Score 2939.5; DB 4; Length 863;  
Best Local Similarity 82.6%; Pred. No. 1.4e-234;  
Matches 568; Conservative 34; Mismatches 39; Indels 47; Gaps 12;  
Qy 1 MRVKGIRKQYQHL-----WRGGTLLGLMLMICSAREKLVVTVYGVVPVWKEATTLFCASD 56  
Db 8 MRV---EKYQHLWRGWRGWTLLGLMLMICSAREKLVVTVYGVVPVWKEATTLFCASD 64  
Qy 57 AKAYTEVHNWATHACVPTDPNPOEIVLENTENFMNMMKNVBMQHEDIISLMDQSLK 116  
Db 65 AKAYTEVHNWATHACVPTDPNPOEIVLVNVTENFMNMMKNVBMQHEDIISLMDQSLK 124  
Qy 117 PCVKLTPLCVTLHCTNKNATNTKSNWKE--MDRGEIKNCSFKVGA--GK----- 163  
Db 125 PCVKLTPLCVSLKCTDLKNDTNTSSSGRMIMKEGEIKNCSFNISTSIKRGVQKEYAFY 184  
Qy 164 -----LINCNTSVITQACPKVFEPIPIHYCAPAGFAILKCNCKKFNKSG 208  
Db 185 KLDIIPIDNDTTSYTLTSCNTSVITQACPKVFEPIPIHYCAPAGFAILKCNCKTNGTG 244  
Qy 209 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEGBGVIRSENFTDNAKTIIIVOLKESVEINCT 268  
Db 245 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEGBGVIRSENFTDNAKTIIIVOLKESVEINCT 304  
Qy 269 RNNNTRKSIIT--GPGRAFATGDIIGDIRAHCNISGEKNNTLKQIVTKLOAOPG-N 325  
Db 305 RNNNTRKSIIRIQRGPGRAFTVIGK--IGNMRAHCNISRAKNWNTLKQIDSKLRQFGNN 363  
Qy 326 KTIIFKQSGGDEIIVTHSFNCGGEFFYCNSTQLFNSTWNN---TIGPNNTNG--TITL 379  
Db 364 KTIIFKQSGGDEIIVTHSFNCGGEFFYCNSTQLFNSTWFSNWTSGNSTEGSDTITL 423  
Qy 380 PCRIKQIINRWQEVGKAMVAPPPIRGQIRCSNITGLLTRDGGKEISNTTEIFRPGGDM 439  
Db 424 PCRIKQIINRWQEVGKAMVAPPISQIRCSNITGLLTRDGGNS--NNESEIFRPGGDM 482  
Qy 440 RDNWRSELYKYKVKVKEPIGLVAPTAKRRVVOREKRAVTLGAMFLGFLGAAGSTMGARSL 499  
Db 483 RDNWRSELYKYKVKVKEPIGLVAPTAKRRVVOREKRAVTLGAMFLGFLGAAGSTMGAASM 542  
Qy 500 TLTVOARQLLSGIVQOONLLRAIEAQHLLQLTVMGIK--LOARVLAVERYLKDQQLLGI 558  
Db 543 TLTVOARQLLSGIVQOONLLRAIEAQHLLQLTVMGIK--LOARVLAVERYLKDQQLLGI 602  
Qy 559 WCGSGKLICTTAVPWNASNSKSLDQIWNMTWMEWERIDNYTNLIYTLIESQNQOEK 618  
Db 603 WCGSGKLICTTAVPWNASNSKSLDQIWNMTWMEWERIDNYTNLIYTLIESQNQOEK 662  
Qy 619 NEQELLELDKWSLWNFIDISKLWYIK 646  
Db 663 NEQELLELDKWSLWNFITNLWYIK 690

Search completed: February 25, 2004, 14:20:10  
Job time : 20.8949 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:18:41 ; Search time 33.7043 Seconds  
(without alignments)  
4047.104 Million cell updates/sec

Title: US-09-891-609A-2  
Perfect score: 3474  
Sequence: 1 MRVKGIRKNYQHLWRGGTLL.....DKWASLNNWFDISKWLWYIK 646

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 segs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/1/pubpaa/FCI\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3474	100.0	646	9	US-09-891-609-2
2	3422	98.5	847	9	US-09-476-242-2
3	3327	95.8	619	9	US-09-891-609-4
4	3279.5	94.4	842	14	US-10-190-435-2
5	3279.5	94.4	842	14	US-10-241-009-2
6	3279.5	94.4	842	14	US-10-190-434B-2
7	3279.5	94.4	842	14	US-10-190-305A-2
8	2937.5	84.6	868	9	US-09-938-406-1
9	2933	84.4	861	14	US-10-026-741-103
10	2922.5	84.1	856	9	US-09-476-242-1
11	2907.5	83.7	856	14	US-10-196-515-11
12	2886	83.1	726	14	US-10-196-515-3
13	2882	83.0	759	14	US-10-196-515-12
14	2862	82.4	643	14	US-10-032-162-13
15	2771	79.8	579	14	US-10-032-162-15

16	2758	79.4	625	14	US-10-032-162-17	Sequence 17, Appl
17	2726	78.5	853	13	US-10-003-035-33	Sequence 33, Appl
18	2726	78.5	853	14	US-10-286-332A-33	Sequence 33, Appl
19	2726	78.5	853	15	US-10-280-915-33	Sequence 33, Appl
20	2700	77.7	860	14	US-10-190-435-6	Sequence 6, Appl
21	2700	77.7	860	14	US-10-241-009-6	Sequence 6, Appl
22	2700	77.7	860	14	US-10-190-434B-6	Sequence 6, Appl
23	2700	77.7	860	14	US-10-190-305A-6	Sequence 6, Appl
24	2667	76.8	870	14	US-10-190-435-147	Sequence 147, App
25	2642.5	76.1	867	14	US-10-190-435-3	Sequence 3, Appl
26	2642.5	76.1	867	14	US-10-190-435-126	Sequence 126, App
27	2642.5	76.1	867	14	US-10-241-009-3	Sequence 3, Appl
28	2642.5	76.1	867	14	US-10-190-434B-3	Sequence 3, Appl
29	2642.5	76.1	867	14	US-10-190-305A-3	Sequence 3, Appl
30	2640.5	76.0	865	14	US-10-190-435-140	Sequence 140, App
31	2635.5	75.9	855	14	US-10-190-435-144	Sequence 144, App
32	2634	75.8	861	14	US-10-190-435-139	Sequence 139, App
33	2632	75.8	858	14	US-10-190-435-150	Sequence 150, App
34	2628	75.6	862	14	US-10-190-435-141	Sequence 141, App
35	2626.5	75.6	855	14	US-10-369-294-11	Sequence 12, Appl
36	2626.5	75.6	855	14	US-10-369-294-12	Sequence 12, Appl
37	2624	75.5	870	14	US-10-190-435-127	Sequence 127, App
38	2621.5	75.5	869	14	US-10-190-435-4	Sequence 4, Appl
39	2621.5	75.5	869	14	US-10-241-009-4	Sequence 4, Appl
40	2621.5	75.5	869	14	US-10-190-434B-4	Sequence 4, Appl
41	2621.5	75.5	869	14	US-10-190-305A-4	Sequence 4, Appl
42	2619.5	75.4	845	14	US-10-190-435-129	Sequence 129, App
43	2619.5	75.4	1101	13	US-10-003-035-53	Sequence 53, Appl
44	2619.5	75.4	1101	14	US-10-286-332A-53	Sequence 53, Appl
45	2619.5	75.4	1101	15	US-10-280-915-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1  
US-09-891-609-2  
; Sequence 2, Application US/09891609  
; Patent No. US20020127238A1  
; GENERAL INFORMATION:  
; APPLICANT: Stratatoc, Leondias  
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR  
; FILE REFERENCE: 2570-1-002N  
; CURRENT APPLICATION NUMBER: US/09/891,609  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/214,608  
; PRIOR FILING DATE: 2000-06-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 2  
; LENGTH: 646  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-891-609-2

Query Match	100.0%;	Score 3474;	DB 9;	Length 646;
Best Local Similarity	100.0%;	Pred. No. 3.5e-310;		
Matches 646;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MRVKGIRKNYQHLWRGGTLLGLMLICSAVEKLWTVVYGVVWKEATTTLCASDAKAY	60	
Db	1	MRVKGIRKNYQHLWRGGTLLGLMLICSAVEKLWTVVYGVVWKEATTTLCASDAKAY	60	
QY	61	DTEHNVNATHACVPTDNPQEIIVLENTENFMKNVVEOMHEDIISLWQSLKPCVK	120	
Db	61	DTEHNVNATHACVPTDNPQEIIVLENTENFMKNVVEOMHEDIISLWQSLKPCVK	120	
QY	121	LTPLCVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSFKVGAGKLINCNTSVITQACP	180	
Db	121	LTPLCVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSFKVGAGKLINCNTSVITQACP	180	
QY	181	FPPIPIHYCAPAGFAILKNDCKKFNCSGPGTNTVSTVQCTHGRPVVSTOLLINGS	240	

Db 181 FEPIPIHYCAPAGFAILKCNKXFGSGCTNVSTVQCTHGIRPVVSTQLLNGSLAERG 240  
Qy 241 VVIRSENFTDNTAKTIIVQLKESVEINCTPNNTKRSITIGPRAFYATGDIIGDIOAH 300  
Db 241 VVIRSENFTDNTAKTIIVQLKESVEINCTPNNTKRSITIGPRAFYATGDIIGDIOAH 300  
Qy 301 CNISGEKWNNTLQKIVTKLQAOFGNKTIVFKOSSGGDDPEIVMHSFNCGGEFFYCNSTOLF 360  
Db 301 CNISGEKWNNTLQKIVTKLQAOFGNKTIVFKOSSGGDDPEIVMHSFNCGGEFFYCNSTOLF 360  
Qy 361 NSTWNTIGPNNTNGTITLPCRKIQIINRWQEVGKAMYAPPIRGQIRCSNITGLLLTRD 420  
Db 361 NSTWNTIGPNNTNGTITLPCRKIQIINRWQEVGKAMYAPPIRGQIRCSNITGLLLTRD 420  
Qy 421 GKKEISNTTTEIPRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVTIG 480  
Db 421 GKKEISNTTTEIPRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVTIG 480  
Qy 481 AMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQOQNLLRAIEAQOHLQLTVWGIKQ 540  
Db 481 AMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQOQNLLRAIEAQOHLQLTVWGIKQ 540  
Qy 541 ARVLAVERYLKDQQLGICGSGKLICTTAVPMNASWSNKSDDQIWNNTWMEWEREIDN 600  
Db 541 ARVLAVERYLKDQQLGICGSGKLICTTAVPMNASWSNKSDDQIWNNTWMEWEREIDN 600  
Qy 601 YTNLIYTLIEESQNOQEKNEOELLELDKWSLWNNWFDISKWLWYIK 646  
Db 601 YTNLIYTLIEESQNOQEKNEOELLELDKWSLWNNWFDISKWLWYIK 646

RESULT 2  
US-09-476-242-2  
; Sequence 2, Application US/09476242  
; Patent No. US20020146683A1  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: HARTOG, Karin  
; APPLICANT: MARTIN, Eric  
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
; FILE REFERENCE: 1605.002  
; CURRENT APPLICATION NUMBER: US/09/476,242  
; CURRENT FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 847  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-476-242-2

Query Match 98.5%; Score 3422; DB 9; Length 847;  
Best Local Similarity 95.4%; Pred. No. 3.1e-305;  
Matches 643; Conservative 0; Mismatches 3; Indels 28; Gaps 2;  
Qy 1 MRVKGIRKKNYQHLWRGGTLLGLMLTCSAVEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
Db 1 MRVKGIRKKNYQHLWRGGTLLGLMLTCSAVEKLWTVVYGVVPWKEATTLFCASDAKAY 60  
Qy 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNWKNMKNMVEQMHEDIISLWDQSLKPCVK 120  
Db 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNWKNMKNMVEQMHEDIISLWDQSLKPCVK 120  
Qy 121 LTPCLVTLHCTNLKNAATNTKSSNWKMDRGEIKNCSFKVTTSTRNKQKEYALFYKLDVV 162  
Db 121 LTPCLVTLHCTNLKNAATNTKSSNWKMDRGEIKNCSFKVTTSTRNKQKEYALFYKLDVV 180  
Qy 163 -----KLINCNTSVITQACPKVSFPIPIHYCAPAGFAILKCNKXFGSGCTNV 213  
Db 181 PIDNDNTSVKILNCNTSVITQACPKVSFPIPIHYCAPAGFAILKCNKXFGSGCTNV 240  
Qy 214 STVQCTHGIRPVVSTQLLNGSLAERGVIIRSENFTDNTAKTIIVQLKESVEINCTPNNT 273

Db 241 STVQCTHGIRPVVSTQLLNGSLAERGVIIRSENFTDNTAKTIIVQLKESVEINCTPNNT 300  
Qy 274 TRKSTITIGPRAFYATGDIIGDIOAHCNISGEKWNNTLQKIVTKLQAOFGNKTIVFKOS 333  
Db 301 TRKSTITIGPRAFYATGDIIGDIOAHCNISGEKWNNTLQKIVTKLQAOFGNKTIVFKOS 360  
Qy 334 SGGDPEIVMHSFNCGGEFFYCNSTOLFNSNTWNTIGPNNTNGTITLPCRKIQIINRWQEV 393  
Db 361 SGGDPEIVMHSFNCGGEFFYCNSTOLFNSNTWNTIGPNNTNGTITLPCRKIQIINRWQEV 420  
Qy 394 GHAMVAPPIRGQIRCSNITGLLLTRDGGKEISNTTEIPRPGGDMRDNRSELYKYKV 453  
Db 421 GHAMVAPPIRGQIRCSNITGLLLTRDGGKEISNTTEIPRPGGDMRDNRSELYKYKV 480  
Qy 454 KIEPLGVAPTAKRRVVQREKRAVTIGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIV 513  
Db 481 KIEPLGVAPTAKRRVVQREKRAVTIGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIV 540  
Qy 514 QOQNLLRAIEAQOHLQLTVWGIK-LQARVLAVERYLKDQQLGICGSGKLICTTAVP 572  
Db 541 QOQNLLRAIEAQOHLQLTVWGIKQOARVLAVERYLKDQQLGICGSGKLICTTAVP 600  
Qy 573 WNASWSNKSDDQIWNNTWMEWEREIDNTNLIYTLIEESQNOQEKNEOELLELDKWSL 632  
Db 601 WNASWSNKSDDQIWNNTWMEWEREIDNTNLIYTLIEESQNOQEKNEOELLELDKWSL 660  
Qy 633 WNWFDISKWLWYIK 646  
Db 661 WNWFDISKWLWYIK 674

RESULT 3  
US-09-891-609-4  
; Sequence 4, Application US/09891609  
; Patent No. US20020127238A1  
; GENERAL INFORMATION:  
; APPLICANT: Stamatos, Leonidas  
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR  
; FILE REFERENCE: 2570-1-002N  
; CURRENT APPLICATION NUMBER: US/09/891,609  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/214,608  
; PRIOR FILING DATE: 2000-06-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 619  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-891-609-4

Query Match 95.8%; Score 3327; DB 9; Length 619;  
Best Local Similarity 100.0%; Pred. No. 1.1e-296;  
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 28 SAVEKLWTVVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIIVLEN 87  
Db 1 SAVEKLWTVVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIIVLEN 60  
Qy 88 VTENFNWKNMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 147  
Db 61 VTENFNWKNMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 120  
Qy 148 DRGEIKNCSFKVAGKILNCNTSVITQACPKVSFPIPIHYCAPAGFAILKCNKXFGS 207  
Db 121 DRGEIKNCSFKVAGKILNCNTSVITQACPKVSFPIPIHYCAPAGFAILKCNKXFGS 180  
Qy 208 GPCTNVSTVQCTHGIRPVVSTQLLNGSLAERGVIIRSENFTDNTAKTIIVQLKESVEINC 267  
Db 181 GPCTNVSTVQCTHGIRPVVSTQLLNGSLAERGVIIRSENFTDNTAKTIIVQLKESVEINC 240  
Qy 268 TRPNNTKRSITIGPRAFYATGDIIGDIOAHCNISGEKWNNTLQKIVTKLQAOFGNKT 327

Db 241 TRPNNTRKSIITIGPGRFYATGDIIGDIRQAHCHNISEKWNNTLKQIVTKLQAOFGNKT 300  
Qy 328 IVFKOSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNTTIGPNNNTNGTITLPCRICKQII 387  
Db 301 IVFKOSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNTTIGPNNNTNGTITLPCRICKQII 360  
Qy 388 NRQOEVGKAMYPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSEL 447  
Db 361 NRQOEVGKAMYPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSEL 420  
Qy 448 YKVKVKIEPLGVAPTKARRVVRQEKRAVTIGAMFLGFLGAAGSTMGARSITLTVQARQ 507  
Db 421 YKVKVKIEPLGVAPTKARRVVRQEKRAVTIGAMFLGFLGAAGSTMGARSITLTVQARQ 480  
Qy 508 LLSGIVQQNNLLRAIEAQHLLQTLVWGIKQARVLAVERYLKDQQLLGIWCGSGKLC 567  
Db 481 LLSGIVQQNNLLRAIEAQHLLQTLVWGIKQARVLAVERYLKDQQLLGIWCGSGKLC 540  
Qy 568 TTAVPWNASWSKSLDQIWNNTWMEWEREIDNYNLIYTLIEESQNOQEKNEQELLED 627  
Db 541 TTAVPWNASWSKSLDQIWNNTWMEWEREIDNYNLIYTLIEESQNOQEKNEQELLED 600  
Qy 628 KWASLWNNWFDISKWLWYIK 646  
Db 601 KWASLWNNWFDISKWLWYIK 619

RESULT 4

US-10-190-435-2  
; Sequence 2, Application US/10190435  
; Publication No. US20030143248A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: BARNETT, Susan W.  
; APPLICANT: LIAN, Ying  
; APPLICANT: ENGELBRECHT, Susan  
; APPLICANT: VAN RENSBURG, Estrelita J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE REFERENCE: PP18133.003 / 2302-18133  
; CURRENT APPLICATION NUMBER: US/10/190,435  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 842  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SF162  
US-10-190-435-2

Query Match 94.4%; Score 3279.5; DB 14; Length 842;  
Best Local Similarity 94.2%; Pred. No. 3.9e-292;  
Matches 619; Conservative 2; Mismatches 7; Indels 29; Gaps 3;  
Qy 19 LLLGLMLMC-SAVEKLLWTVVYGVVPWKEATTTLCASDAKAYDEVHNVWATHACVPTD 77  
Db 13 LLCGAVFVSPSAVEKLLWTVVYGVVPWKEATTTLCASDAKAYDEVHNVWATHACVPTD 72  
Qy 78 PNQBEIVLNVNTENFMNWKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAT 137  
Db 73 PNQBEIVLNVNTENFMNWKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAT 132  
Qy 138 NTKSSNWKEMDRGEIKNCSFKVAG-----KLNINCTS 170  
Db 133 NTKSSNWKEMDRGEIKNCSFKVTTIRNMOKKEYALFYKLDVVPIDNDNTSYKLNINCTS 192  
Qy 171 VITOACPKVSFPIPIHYCAPAGFAILKCNCKKFGSGPCTNVSTVQCTHGRIPVVSTQL 230  
Db 193 VITOACPKVSFPIPIHYCAPAGFAILKCNCKKFGSGPCTNVSTVQCTHGRIPVVSTQL 252  
Qy 231 LLNGSLAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRFYATG 290

Db 253 LLNGSLAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRFYATG 312  
Qy 291 DIIGDIRQAHCHNISEKWNNTLKQIVTKLQAOFGNKTIVFKOSSGGDPEIVMHSFNCGG 350  
Db 313 DIIGDIRQAHCHNISEKWNNTLKQIVTKLQAOFGNKTIVFKOSSGGDPEIVMHSFNCGG 372  
Qy 351 FFYCNSTQLFNSWNTTIGPNNNTNGTITLPCRICKQIINRWQEVGKAMYPPIRGQIRCS 410  
Db 373 FFYCNSTQLFNSWNTTIGPNNNTNGTITLPCRICKQIINRWQEVGKAMYPPIRGQIRCS 432  
Qy 411 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELKYKVKVKIEPLGVAPTKARRV 470  
Db 433 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELKYKVKVKIEPLGVAPTKARRV 492  
Qy 471 QREKRAVTIGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQQNNLLRAIEAQHLL 530  
Db 493 QREKRAVTIGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQQNNLLRAIEAQHLL 552  
Qy 531 QLTWVGIK-LQARVLAVERYLKDQQLLGIWCGSGKLCITTA VPMNASWSKSLDQIWNNT 589  
Db 553 QLTWVGIKQARVLAVERYLKDQQLLGIWCGSGKLCITTA VPMNASWSKSLDQIWNNT 612  
Qy 590 TWMEWEREIDNYNLIYTLIEESQNOQEKNEQELLEDKWSLWNNWFDISKWLWYIK 646  
Db 613 TWMEWEREIDNYNLIYTLIEESQNOQEKNEQELLEDKWSLWNNWFDISKWLWYIK 669

RESULT 5

US-10-241-009-2  
; Sequence 2, Application US/10241009  
; Publication No. US20030170614A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: BARNETT, Susan  
; APPLICANT: LIAN, Ying  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B  
; FILE REFERENCE: 2300-1621.21  
; CURRENT APPLICATION NUMBER: US/10/241,009  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 842  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SF162  
US-10-241-009-2

Query Match 94.4%; Score 3279.5; DB 14; Length 842;  
Best Local Similarity 94.2%; Pred. No. 3.9e-292;  
Matches 619; Conservative 2; Mismatches 7; Indels 29; Gaps 3;  
Qy 19 LLLGLMLMC-SAVEKLLWTVVYGVVPWKEATTTLCASDAKAYDEVHNVWATHACVPTD 77  
Db 13 LLCGAVFVSPSAVEKLLWTVVYGVVPWKEATTTLCASDAKAYDEVHNVWATHACVPTD 72  
Qy 78 PNQBEIVLNVNTENFMNWKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAT 137  
Db 73 PNQBEIVLNVNTENFMNWKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAT 132  
Qy 138 NTKSSNWKEMDRGEIKNCSFKVAG-----KLNINCTS 170  
Db 133 NTKSSNWKEMDRGEIKNCSFKVTTIRNMOKKEYALFYKLDVVPIDNDNTSYKLNINCTS 192  
Qy 171 VITOACPKVSFPIPIHYCAPAGFAILKCNCKKFGSGPCTNVSTVQCTHGRIPVVSTQL 230  
Db 193 VITOACPKVSFPIPIHYCAPAGFAILKCNCKKFGSGPCTNVSTVQCTHGRIPVVSTQL 252  
Qy 231 LLNGSLAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRFYATG 290

Db 253 LLNGSLAEBGVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITIGPRAFYATG 312  
QY 291 DIIGDIRQAHCHNISGEKNWNTLKQIVTKLQAOFGNKTIVFKOSSGDDPEIVMHSFCGGE 350  
Db 313 DIIGDIRQAHCHNISGEKNWNTLKQIVTKLQAOFGNKTIVFKOSSGDDPEIVMHSFCGGE 372  
QY 351 FFYCNSTQLFNSNTWNTTIGPNNNTGTITLPCRIKQIINRWQVSGKAMVAPPPIRGQIRCSS 410  
Db 373 FFCNSTQLFNSNTWNTTIGPNNNTGTITLPCRIKQIINRWQVSGKAMVAPPPIRGQIRCSS 432  
QY 411 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVV 470  
Db 433 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVV 492  
QY 471 QREKRAVTLGAMFLGFLGAAGSTMGARSILTLTVQARQLLSGIVQOQNLLRAIEAQOHL 530  
Db 493 QREKRAVTLGAMFLGFLGAAGSTMGARSILTLTVQARQLLSGIVQOQNLLRAIEAQOHL 552  
QY 531 QLTVMGIK-LQARVLAVERYLKDOQLLGIWCGSGKLICTTAVPMNASWNSKSLDQIWNM 589  
Db 553 QLTVMGIKQ-LQARVLAVERYLKDOQLLGIWCGSGKLICTTAVPMNASWNSKSLDQIWNM 612  
QY 590 TWMEREIDNVTNLIYTLIESQNOEKNEQELLELDKWASLWNNWFDISKWLWYIK 646  
Db 613 TWMEREIDNVTNLIYTLIESQNOEKNEQELLELDKWASLWNNWFDISKWLWYIK 669

## RESULT 6

US-10-190-434B-2

; Sequence 2, Application US/10190434B  
; Publication No. US20030194800A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B

; FILE REFERENCE: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

; CURRENT FILING DATE: 2002-07-05

; CURRENT APPLICATION NUMBER: US/10/190,434B

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SF162

US-10-190-434B-2

Query Match 94.4%; Score 3279.5; DB 14; Length 842;  
Best Local Similarity 94.2%; Pred. No. 3.9e-292;  
Matches 619; Conservative 2; Mismatches 7; Indels 29; Gaps 3;

QY 19 LLLGLMLIC-SAVEKLVTVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTD 77  
Db 13 LLCGAVFVSPSAVEKLVTVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTD 72  
QY 78 PNPQEIIVLENTVFNNMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNA 137  
Db 73 PNPQEIIVLENTVFNNMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNA 132  
QY 138 NTKSSNWKMDREIGKCSFKVAG-----KLINCNTS 170  
Db 133 NTKSSNWKMDREIGKCSFKVTTTSIRNMQKEYALFYKLDVVVPIDNDNTSYKLINCNTS 192  
QY 171 VITOACPKVSPFPIPIHYCAPAGFAILKCNDDKFFNGSGPCTNVSTVQCTHGIRPVVSTQL 230  
Db 193 VITOACPKVSPFPIPIHYCAPAGFAILKCNDDKFFNGSGPCTNVSTVQCTHGIRPVVSTQL 252  
QY 231 LLNGSLAEBGVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITIGPRAFYATG 290  
Db 253 LLNGSLAEBGVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITIGPRAFYATG 312

QY 291 DIIGDIRQAHCHNISGEKNWNTLKQIVTKLQAOFGNKTIVFKOSSGDDPEIVMHSFCGGE 350  
Db 313 DIIGDIRQAHCHNISGEKNWNTLKQIVTKLQAOFGNKTIVFKOSSGDDPEIVMHSFCGGE 372  
QY 351 FFYCNSTQLFNSNTWNTTIGPNNNTGTITLPCRIKQIINRWQVSGKAMVAPPPIRGQIRCSS 410  
Db 373 FFCNSTQLFNSNTWNTTIGPNNNTGTITLPCRIKQIINRWQVSGKAMVAPPPIRGQIRCSS 432  
QY 411 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVV 470  
Db 433 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVV 492  
QY 471 QREKRAVTLGAMFLGFLGAAGSTMGARSILTLTVQARQLLSGIVQOQNLLRAIEAQOHL 530  
Db 493 QREKRAVTLGAMFLGFLGAAGSTMGARSILTLTVQARQLLSGIVQOQNLLRAIEAQOHL 552  
QY 531 QLTVMGIK-LQARVLAVERYLKDOQLLGIWCGSGKLICTTAVPMNASWNSKSLDQIWNM 589  
Db 553 QLTVMGIKQ-LQARVLAVERYLKDOQLLGIWCGSGKLICTTAVPMNASWNSKSLDQIWNM 612  
QY 590 TWMEREIDNVTNLIYTLIESQNOEKNEQELLELDKWASLWNNWFDISKWLWYIK 646  
Db 613 TWMEREIDNVTNLIYTLIESQNOEKNEQELLELDKWASLWNNWFDISKWLWYIK 669

## RESULT 7

US-10-190-305A-2

; Sequence 2, Application US/10190305A

; Publication No. US20030198621A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR

; FILE REFERENCE: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

; CURRENT FILING DATE: 2002-07-05

; CURRENT APPLICATION NUMBER: US/10/190,305A

; NUMBER OF SEQ ID NOS: 93

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SF162

US-10-190-305A-2

Query Match 94.4%; Score 3279.5; DB 14; Length 842;  
Best Local Similarity 94.2%; Pred. No. 3.9e-292;  
Matches 619; Conservative 2; Mismatches 7; Indels 29; Gaps 3;

QY 19 LLLGLMLIC-SAVEKLVTVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTD 77  
Db 13 LLCGAVFVSPSAVEKLVTVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTD 72  
QY 78 PNPQEIIVLENTVFNNMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNA 137  
Db 73 PNPQEIIVLENTVFNNMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNA 132  
QY 138 NTKSSNWKMDREIGKCSFKVAG-----KLINCNTS 170  
Db 133 NTKSSNWKMDREIGKCSFKVTTTSIRNMQKEYALFYKLDVVVPIDNDNTSYKLINCNTS 192  
QY 171 VITOACPKVSPFPIPIHYCAPAGFAILKCNDDKFFNGSGPCTNVSTVQCTHGIRPVVSTQL 230  
Db 193 VITOACPKVSPFPIPIHYCAPAGFAILKCNDDKFFNGSGPCTNVSTVQCTHGIRPVVSTQL 252  
QY 231 LLNGSLAEBGVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITIGPRAFYATG 290  
Db 253 LLNGSLAEBGVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITIGPRAFYATG 312

QY 291 DIIGDROAHCHNISGEKWNNTLKQIVTKLQAOFGNKTIVFKOSSGGDPBIVMHSFNCGGE 350  
DB 313 DIIGDROAHCHNISGEKWNNTLKQIVTKLQAOFGNKTIVFKOSSGGDPBIVMHSFNCGGE 372  
QY 351 FFCVNSTQLPNSNTWNTGNTITLPCRILQIINRWQEVGKAMVAPPVIRGOIRCS 410  
DB 373 FFCVNSTQLPNSNTWNTGNTITLPCRILQIINRWQEVGKAMVAPPVIRGOIRCS 432  
QY 411 NITGLLLTRDGGKEISNTTIEIFRPGGDMRDNRSELYKYKVKIPLGVAPTKARRVY 470  
DB 433 NITGLLLTRDGGKEISNTTIEIFRPGGDMRDNRSELYKYKVKIPLGVAPTKARRVY 492  
QY 471 OREKRAVTIGAMFLGFLGAGSTMARSITLTVQARQLLSGIVQQNNLLRAIEAQHLL 530  
DB 493 OREKRAVTIGAMFLGFLGAGSTMARSITLTVQARQLLSGIVQQNNLLRAIEAQHLL 552  
QY 531 QLTVMGK-LOARVLAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDQIWNM 589  
DB 553 QLTVMGKLOARVLAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDQIWNM 612  
QY 590 TWMEWEREIDNTNLIYTLIESQNOEKNEQELLELDKWSLWNNFDISKWLWYIK 646  
DB 613 TWMEWEREIDNTNLIYTLIESQNOEKNEQELLELDKWSLWNNFDISKWLWYIK 669

RESULT 8  
US-09-938-406-1  
; Sequence 1, Application US/09938406  
; Patent No. US20020155120A1  
; GENERAL INFORMATION:  
; APPLICANT: Lowell, George  
; APPLICANT: Vancott, Thomas  
; APPLICANT: Birx, Deborah  
; TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR  
; TITLE OF INVENTION: INDUCING MUCOSAL IMMUNITY  
; FILE REFERENCE: 40646-20002.10  
; CURRENT APPLICATION NUMBER: US/09/938,406  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 09/214,701  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: PCT/US 97/12253  
; PRIOR FILING DATE: 1997-07-10  
; PRIOR APPLICATION NUMBER: US 60/021,687  
; PRIOR FILING DATE: 1996-07-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Virus HIV-1  
US-09-938-406-1

Query Match 84.6%; Score 2937.5; DB 9; Length 868;  
Best Local Similarity 79.5%; Pred. No. 1.1e-260;  
Matches 556; Conservative 37; Mismatches 47; Indels 59; Gaps 8;

QY 1 MRVKGIRKYNQHLWRGGTLLGLMLICSAVEKLWTVVYGVVPWKAEATTLFCASDAKAY 60  
DB 3 MRAGIRKNCQHLWRGWTLLGLMLICSAANLWTVVYGVVPWKAEATTLFCASDAKAY 62  
QY 61 DTEHNVWATHACVPTDNPQBIENVNTFNFMKNNMVEQMHEDIISLWQSLKPCVK 120  
DB 63 DTEHNVWATHACVPTDNPQBIENVNTFNFMKNNMVEQMHEDIISLWQSLKPCVK 122  
QY 121 LTPCLVTLHCT--NLKNTNTKSSN-----WKMDRGEIKNCSFKVAG----- 162  
DB 123 LTPCLVTLNCTDLNTNTTITLSIIVWQEGKGMENCSFNITTSIRDKVQREYALF 182  
QY 163 -----KLINCNTSVITQACPKVSEPIPIHYCAPAGFAILKCNCK 202  
DB 183 YKLDVEPIDDKNTNTKYRLINCNTSVITQACPKVSEPIPIHYCTPTGFALLKCNCK 242  
QY 203 KFGSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEVGVVIRSENFDTNNAKTIIVQLKES 262

DB 243 KFGSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEVGVVIRSENFDTNNAKTIIVQLKES 302  
QY 263 VEINCTRPNNTRKSIITIGPGRFYATGDIIGDROAHCHNISGEKWNNTLKQIVTKLQAO 322  
DB 303 VEINCTRPNNTRKSIITIGPGRVTTTGEILGNIROAHCHNISRAQWNNTLQOQIATLRQ 362  
QY 323 FGNKTIIVFKOSSGGDPBIVMHSFNCGGEFFYCNSTQLFNSTWNNTIGPNNTNGT----- 376  
DB 363 FGNKTIIVFKOSSGGDPBIVMHSFNCGGEFFYCNSTQLFNSANVT-----SNGTWSVTRK 417  
QY 377 -----ITLPCRILQIINRWQEVGKAMVAPPVIRGOIRCSNITGLLLTRDGGKEISNTT 429  
DB 418 QKDTGDIITLPCRILQIINRWQVWKAMVAPPVIRGOIRCSNITGLLLTRDGGGE-NQTT 476  
QY 430 EIFRPGGDMRDNRSELYKYKVKIPLGVAPTKARRVVOREKRAV-TLGAMFLGFLG 488  
DB 477 EIFRPGGDMRDNRSELYKYKVKIPLGVAPTKARRVVOREKRAVGLAMFLGFLG 536  
QY 489 AAGSTMARSITLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTVMGK-LOARVLAVE 547  
DB 537 AAGSTMARSITLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTVMGKLOARVLAVE 596  
QY 548 RYLDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDQIWNNTWMEWEREIDNTNLIY 607  
DB 597 RYLDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDQIWNNTWMEWEREIDNTNLIY 656  
QY 608 LIEESQNOEKNEQELLELDKWSLWNNFDISKWLWYIK 646  
DB 657 LIEESQNOEKNEQELLELDKWSLWNNFDISKWLWYIK 695

RESULT 9  
US-10-026-741-103  
; Sequence 103, Application US/10026741  
; Publication No. US20030049604A1  
; GENERAL INFORMATION:  
; APPLICANT: CHARNEAU, PIERRE  
; BORMAN, ANDREW  
; QUILLENT, CAROLINE  
; GUETARD, DENISE  
; MONTAGNIER, LUC  
; DONJON DE SAINT-MARTIN, JACOBELINE  
; COHEN, JAUQUES  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; Dunner, L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/026,741  
; FILING DATE: 27-Dec-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/817,441  
; FILING DATE: 31-AUG-1998  
; APPLICATION NUMBER: PCT/FR 95/01391  
; FILING DATE: 20-OCT-1995  
; APPLICATION NUMBER: FR 9412554  
; FILING DATE: 20-OCT-1994  
; APPLICATION NUMBER: FR 9502526  
; FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03260.6005-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 861 amino acids  
TYPE: amino acid  
STRAINEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
US-10-026-741-103

Query Match 84.4%; Score 2933; DB 14; Length 861;  
Best Local Similarity 81.8%; Pred. No. 2.8e-260;  
Matches 567; Conservative 33; Mismatches 41; Indels 52; Gaps 12;  
QY 1 MRVKGIRKNYQHLR-----GGTLLGLMLMICS AVEKLVVTVYGVVPVWKEATTTFLFCASD 56  
DB 1 MRVK---EKYQHLRWGRGWGTMLLGLMICS ATEKLVVTVYGVVPVWKEATTTFLFCASD 57  
QY 57 AKAYDEVHNWATHACVPTDPNPOEIVLENTFNFMKNMKNVQEHEDIIISLWDSQSLK 116  
DB 58 AKAYDEVHNWATHACVPTDPNPOEVVLVNTFNFMKNMKNVQEHEDIIISLWDSQSLK 117  
QY 117 PCVKLTPLCVLTHCTNLKNATNTKSSNWKE-----MDRGEIKNCSFKVGA---GK----163  
DB 118 PCVKLTPLCVSLKCTDLGNATNTSSNTSSSGEMMEKEIKNCSFNISTSRGKVQKE 177  
QY 164 -----LINCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNDDK 203  
DB 178 YAFFYKLDIIPDNDDTTSYTLTSCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNKNT 237  
QY 204 FNGSGCTNVSTVQCTHGIRPVVSTOLLNGSLAEVGVVIRSENFDTNAKTIIIVQLKESV 263  
DB 238 FNGTGTNVSTVQCTHGIRPVVSTOLLNGSLAEVEVIRSENFDTNAKTIIIVQLNQS 297  
QY 264 EINCTRPNNTRKSIITI--GPGRAFYATGDIIGDIRQAHNCISGEKNWNTLKQIVTKLOA 321  
DB 298 EINCTRPNNTRKSIIRIQRGPGRAFYTIK-IGNMRQAHNCISRAKWNATLQIASKLRE 356  
QY 322 QFG-NKTIIVFKQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSWN--TIGPNTNG-375  
DB 357 QFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSWTFNSWSTEGSNNTGS 416  
QY 376 -TITLPCRIRKQIINRWQEVGKAMYAPPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRP 434  
DB 417 DTITLPCRIRKQFINMWQEVGKAMYAPPISGQIRCSSNITGLLLTRDGGNN-ANGSEIFRP 475  
QY 435 GGGDMRDNRSELYKYKVKVIEPLGVAPTAKARRVQREKRAVTLGAMFLGFLGAAGSTM 494  
DB 476 GGGDMRDNRSELYKYKVKVIEPLGVAPTAKARRVQREKRAVGIGALFLGFLGAAGSTM 535  
QY 495 GARSLLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTVMGIK-LQARVLAVERYLKQQLGI 553  
DB 536 GARSMTLTQARQLLSGIVQOQNLLRAIEAQHLLQLTVMGIKQLQARILAVERYLKQ 595  
QY 554 QLIGIWCSCGLICTTAVPWNASWSKSLDOIWNMTWMEWEIREIDNYTNLIYTLIEESQ 613  
DB 596 QLIGIWCSCGLICTTAVPWNASWSKSLDOIWNMTWMEWEIREIDNYTNLIYTLIEESQ 655  
QY 614 NQEKNEQELLELDKWSLWNPFDISKWLWYIK 646  
DB 656 NQEKNEQELLELDKWSLWNPFDISKWLWYIK 688

RESULT 10

US-09-476-242-1

; Sequence 1, Application US/09476242

Patent No. US20020146683A1  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: HARTOG, Karin  
APPLICANT: MARTIN, Eric  
TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
FILE REFERENCE: 1605.002  
CURRENT APPLICATION NUMBER: US/09/476,242  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1  
LENGTH: 856  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-476-242-1

Query Match 84.1%; Score 2922.5; DB 9; Length 856;  
Best Local Similarity 81.8%; Pred. No. 2.6e-259;  
Matches 563; Conservative 35; Mismatches 43; Indels 47; Gaps 11;  
QY 1 MRVKGIRKNYQHL-----WRGGTLLGLMLMICS AVEKLVVTVYGVVPVWKEATTTFLFCASD 56  
DB 1 MRVK---EKYQHLRWGRGWGTMLLGLMICS ATEKLVVTVYGVVPVWKEATTTFLFCASD 57  
QY 57 AKAYDEVHNWATHACVPTDPNPOEIVLENTFNFMKNMKNVQEHEDIIISLWDSQSLK 116  
DB 58 AKAYDEVHNWATHACVPTDPNPOEVVLVNTFNFMKNMKNVQEHEDIIISLWDSQSLK 117  
QY 117 PCVKLTPLCVLTHCTNLKNATNTKSSNWKE-MDRGEIKNCSFKVGA-----162  
DB 118 PCVKLTPLCVSLKCTDLKNDTNTSSSGRMIMEKEIKNCSFNISTSRGKVQKEAFFY 177  
QY 163 -----KLINCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNDDK 208  
DB 178 KLDIIPDNDDTTSYTLTSCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNKNTFNGT 237  
QY 209 PCTNVSTVQCTHGIRPVVSTOLLNGSLAEVGVVIRSENFDTNAKTIIIVQLKESV 268  
DB 238 PCTNVSTVQCTHGIRPVVSTOLLNGSLAEVEVIRSENFDTNAKTIIIVQLNQS 297  
QY 269 RPNNTNRKSIITI--GPGRAFYATGDIIGDIRQAHNCISGEKNWNTLKQIVTKLOA 325  
DB 298 RPNNTNRKSIIRIQRGPGRAFYTIK-IGNMRQAHNCISRAKWNATLQIASKLREQGNN 356  
QY 326 KTIIVFKQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSWN--TIGPNTNG--TITL 379  
DB 357 KTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSWTFNSWSTEGSNNTGS 416  
QY 380 PCRIRKQIINRWQEVGKAMYAPPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRP 439  
DB 417 PCRIRKQIINMWQEVGKAMYAPPISGQIRCSSNITGLLLTRDGGNS-NNSEIFRP 475  
QY 440 RDNWRSELYKYKVKVIEPLGVAPTAKARRVQREKRAVTLGAMFLGFLGAAGSTM 499  
DB 476 RDNWRSELYKYKVKVIEPLGVAPTAKARRVQREKRAVGIGALFLGFLGAAGSTM 535  
QY 500 TLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTVMGIK-LQARVLAVERYLKQQLGI 558  
DB 536 TLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTVMGIKQLQARILAVERYLKQQLGI 595  
QY 559 WCSGKGLICTTAVPWNASWSKSLDOIWNMTWMEWEIREIDNYTNLIYTLIEESQ 618  
DB 596 WCSGKGLICTTAVPWNASWSKSLDOIWNMTWMEWEIREIDNYTNLIYTLIEESQ 655  
QY 619 NEQELLELDKWSLWNPFDISKWLWYIK 646  
DB 656 NEQELLELDKWSLWNPFDISKWLWYIK 683

RESULT 11

US-10-196-515-11

; Sequence 11, Application US/10196515

Publication No. US20030091594A1  
GENERAL INFORMATION:  
APPLICANT: HOXIE, James A.  
APPLICANT: LABRANCHE, Celia C.  
APPLICANT: DOMS, Robert W.  
APPLICANT: HOFFMAN, Trevor L.  
TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
TITLE OF INVENTION: THERAPEUTICS  
FILE REFERENCE: Hoxie 9596-104U1 (0282)  
CURRENT APPLICATION NUMBER: US/10/196,515  
CURRENT FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: US/09/337,387  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: US 09/317,556  
PRIOR FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 856  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-10-196-515-11

Query Match 83.7%; Score 2907.5; DB 14; Length 856;  
Best Local Similarity 81.7%; Pred. No. 6.2e-258;  
Matches 562; Conservative 37; Mismatches 42; Indels 47; Gaps 12;

QY 1 MRVKGIRKYNQHL---WRGGTLLGLMLICSAVEKLWTVVYGVVWKEATTTILFCASD 56  
DB 1 MRVK---EKYQHLRWGRWGTMLGLMLICNATEKLWTVVYGVVWKEATTTILFCASD 57  
QY 57 AKAYDEVHNVWATHACVPTDPNPQEIIVLENTFNFMNKNMVEQMHEDIISLWDQSLK 116  
DB 58 AKAYDEVHNVWATHACVPTDPNPQEVVLVNTFNFMNKNMVEQMHEDIISLWDQSLK 117  
QY 117 PCVKLTPLCVLHCTNLKNATKSNWKE-MDRGEIKNCSFKVGA--GK-----163  
DB 118 PCVKLTPLCVLSLCTDLKNDTNTSSSGRMIMEKEIKNCSFNISTIRKVKQYAFY 177  
QY 164 -----LINCNTSVITQACPVSFEPIPIHYCAPAGFALLKCNKDKFNGSG 208  
DB 178 KLDIIPDNDTYSLSNTSVITQACPVSFEPIPIHYCAPAGFALLKCNKDKFNGSG 237  
QY 209 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEGBGVIRSENFTDNAKTIIIVQLKESVEINCT 268  
DB 238 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEGBGVIRSENFTDNAKTIIIVQLKESVEINCT 297  
QY 269 RNNNTRKSIITI--GPGRAFATGDIIGDIRQAHNCISGEKWNNTLKQIVTKLQAFG-N 325  
DB 298 RNNNTRKRIRIQRGPGRAFVTIGK-IGNMROAHCNISRAKWNNTLKQIDSKLREQFGNN 356  
QY 326 KTIIVKQSSGGDPEIVMHSFNCGGEFFYCNSTOLFNSWTNN---TIGPNNTNG--TITL 379  
DB 357 KTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTOLFNSWTNNSTWSTEGSNTGSDTITL 416  
QY 380 PCRIKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGGDM 439  
DB 417 PCRIKQIINWQKVGKAMYPPIRGQIRCSNITGLLLTRDGGNS--NNESEIFRPGGGDM 475  
QY 440 RNNRSELYKVKVKEPIEPLGVAPTKAKRRVQREKRAVTLGAMFLGFLGAAGSTWARS 499  
DB 476 RNNRSELYKVKVKEPIEPLGVAPTKAKRRVQREKRAVIGALFLGFLGAAGSTWGAASM 535  
QY 500 TLTVOARQLLSGIVQOONLLRAIEAQHLLQLTWVGIK-LOARVLAVERYLKDDQLLGI 558  
DB 536 TLTVOARQLLSGIVQOONLLRAIEAQHLLQLTWVGIKLOARVLAVERYLKDDQLLGI 595  
QY 559 WCGSGKLICTTAVPNWASNSKSLDQIWNNTWMEWERIDVNTLYLIBESNQOQEK 618  
DB 596 WCGSGKLICTTAVPNWASNSKSLQIWNHTWMEWDREINNTLYLIBESNQOQEK 655  
QY 619 NQOELLELDKWSLWNFIDISKWLWYIK 646  
|||||

Db 656 NQOELLELDKWSLWNFNTWNLWYIK 683  
RESULT 12  
US-10-196-515-3  
Sequence 3, Application US/10196515  
Publication No. US20030091594A1  
GENERAL INFORMATION:  
APPLICANT: HOXIE, James A.  
APPLICANT: LABRANCHE, Celia C.  
APPLICANT: DOMS, Robert W.  
APPLICANT: HOFFMAN, Trevor L.  
TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
TITLE OF INVENTION: THERAPEUTICS  
FILE REFERENCE: Hoxie 9596-104U1 (0282)  
CURRENT APPLICATION NUMBER: US/10/196,515  
CURRENT FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: US/09/337,387  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: US 09/317,556  
PRIOR FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 726  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-10-196-515-3

Query Match 83.1%; Score 2886; DB 14; Length 726;  
Best Local Similarity 81.0%; Pred. No. 4.6e-256;  
Matches 554; Conservative 39; Mismatches 47; Indels 44; Gaps 11;

QY 1 MRVKGIRKYNQHL---WRGGTLLGLMLICSAVEKLWTVVYGVVWKEATTTILFCASD 56  
DB 1 MRVK---EKYQHLRWGRWGTMLGLMLICNATEKLWTVVYGVVWKEATTTILFCASD 57  
QY 57 AKAYDEVHNVWATHACVPTDPNPQEIIVLENTFNFMNKNMVEQMHEDIISLWDQSLK 116  
DB 58 AKAYDEVHNVWATHACVPTDPNPQEVVLVNTFNFMNKNMVEQMHEDIISLWDQSLK 117  
QY 117 PCVKLTPLCVLHCTNLKNATKSNWKE-MDRGEIKNCSFKVGA-----163  
DB 118 PCVKLTPLCVLSLCTDLKNDTNTSSSGRMIMEKEIKNCSFNISTIRKVKQYAFY 177  
QY 164 -----LINCNTSVITQACPVSFEPIPIHYCAPAGFALLKCNKDKFNGSG 208  
DB 178 KLDIIPDNDTYSLSNTSVITQACPVSFEPIPIHYCAPAGFALLKCNKDKFNGSG 237  
QY 209 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEGBGVIRSENFTDNAKTIIIVQLKESVEINCT 268  
DB 238 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEGBGVIRSENFTDNAKTIIIVQLKESVEINCT 297  
QY 269 RNNNTRKSIITI--GPGRAFATGDIIGDIRQAHNCISGEKWNNTLKQIVTKLQAFG-N 325  
DB 298 KNNNTRKRIRIHRGPGRAFVTIGK-IGNMROAHCNISRAKWNNTLKQIASKLREQFGNN 356  
QY 326 KTIIVKQSSGGDPEIVMHSFNCGGEFFYCNSTOLFNSWTNNNTIGPNNTNG--TITLPCR 383  
DB 357 KTIIFKQSSGGDPEIVTHSFNCGGEFFYCKSTOLFNSWT-STKGSNTTSGSDTITLPCR 415  
QY 384 KQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGGDMRDNW 443  
DB 416 KQVINWQEVGKAMYPPIRGQIRCSNITGLLLTRDGGNS--NNESEIFRPGGGDMRDNW 474  
QY 444 RSELYKVKVKEPIEPLGVAPTKAKRRVQREKRAVTLGAMFLGFLGAAGSTWARSJTLTV 503  
DB 475 RSELYKVKVKEPIEPLGVAPTKAKRRVQREKRAVIGALFLGFLGAAGSTWGAASMTLV 534  
QY 504 QARQLLSGIVQOONLLRAIEAQHLLQLTWVGIK-LOARVLAVERYLKDDQLLGTWGS 562  
DB 535 QARQLLSGIVQOONLLRAIEAQHLLQLTWVGIKLOARVLAVERYLKDDQLLGTWGS 594  
|||||





Db 420 GGIN-ENGTEIFRPGGDMRDNRSEFYKYKVVVIEPLGVAPTCKRRVVQREKRAVGIG 478  
Qy 481 AMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQQONNLLRAIEAQOHLQLTVMGK-L 539  
Db 479 AVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQQONNLLRAIEAQOHLQLTVMGKQL 538  
Qy 540 QARVLAVERYLKDQQLLIGWCSGKLICTTAVPWNASWNSKSLDQIWNNTWMEWEREID 599  
Db 539 QARVLAVERYLGDQQLLIGWCSGKLICTTAVPWNASWNSKSLDRINNNNTWMEWEREID 598  
Qy 600 NYTNLIYTLIESONQOEKNEQELLELDKWASLWNNWFDISKWLWY 644  
Db 599 NYTSEIYTLIESONQOEKNEQELLELDKWASLWNNWFDITNWLWY 643

RESULT 15  
US-10-032-162-15  
; Sequence 15, Application US/10032162  
; Publication No. US20030052839A1  
; GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADDON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331az  
; CURRENT APPLICATION NUMBER: US/10/032,162  
; PRIOR FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: 09/602,864  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-032-162-15

Query Match 79.8%; Score 2771; DB 14; Length 579;  
Best Local Similarity 85.6%; Pred. No. 1.2e-245;  
Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;  
Qy 30 VEKLWTVVYGVVPVWKEATTLFCASDAKAYDEVHNWATHACVPTDPNPOEIVLENTV 89  
Db 1 VEKLWTVVYGVVPVWKEATTLFCASDAKAYDEVHNWATHACVPTDPNPOEIVLENTV 60  
Qy 90 ENFNWKNMNVQOMHEDIISLWDQSLKPCVKLTPLCVTLCTNLKKNATNTKSSNWKEMDR 149  
Db 61 EHFNMWKNMNVQOMHEDIISLWDQSLKPCVKLTPLC----- 96  
Qy 150 GEIKNCSFKVGAGKLINCNTSVITQACPKVSEPIPIHYCAPAGFALLKCNCKENGSGP 209  
Db 97 -----GAG-----CDSVITQACPKISPEPIPIHYCAPAGFALLKCNCKTENGKGP 142  
Qy 210 CTNVSTVQCTHGIRPVSTQLLNGSLAEEGVIVRSENFDTNAKTIIVOLKESVEINCTR 269  
Db 143 CKNVSTVQCTHGIRPVSTQLLNGSLAEVEVIRSDNFTNNAKTIIVOLKESVEINCTR 202  
Qy 270 PNNNTRKSTIGPGRAFATGDIIGDROAHNCNIGSEKNNTLKQIVTKLQAFQGNKTIV 329  
Db 203 PNNNTRKSTIHGPGAFYTTGIIIGDROAHNCNISRAKNWDTLKQIVIKLREQFENKTIV 262  
Qy 330 FKQSSGDDPEIWMHFNCGGEFFPCNSTOLFNSW-NNTIGPNTNG-TITLPCRKQII 387  
Db 263 FNHSSGDDPEIWMHFNCGGEFFPCNSTOLFNSWNNNTGSGNTEGNTITLPCRKQII 322  
Qy 388 NRWQEVGKAYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSEL 447  
Db 323 NMWQEVGKAYAPPPIRGQIRCSNITGLLLTRDGGIN-ENGTEIFRPGGDMRDNRSEL 381  
Qy 448 YKYKVVVIEPLGVAPTCKRRVVQREKRAVLTGAMFLGFLGAAGSTMGARSITLTVQARQ 507

Db 382 YKYKVVVIEPLGVAPTCKRRVVQREKRAVLTGAMFLGFLGAAGSTMGAASMTLTVQARL 441  
Qy 508 LLSGIVQQONNLLRAIEAQOHLQLTVMGK-LQARVLAVERYLKDQQLLIGWCSGKLI 566  
Db 442 LLSGIVQQONNLLRAIEAQOHLQLTVMGKQLQARVLAVERYLGDQQLLIGWCSGKLI 501  
Qy 567 CTTAVPWNASWNSKSLDQIWNNTWMEWEREIDNTYTNLIYTLIESONQOEKNEQELLEL 626  
Db 502 CCTAVPWNASWNSKSLDRINNNNTWMEWEREIDNTYTSIYTLIESONQOEKNEQELLEL 561  
Qy 627 DKWASLWNNWFDISKWLWY 644  
Db 562 DKWASLWNNWFDITNWLWY 579

Search completed: February 25, 2004, 14:28:09  
Job time : 36.7043 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:12:55 ; Search time 18.3842 Seconds  
(without alignments)  
3380.063 Million cell updates/sec

Title: US-09-891-609A-2  
Perfect score: 3474  
Sequence: 1 MRVKGIRKYNQHLWRGGTLL.....DKWASLWNWFDISKWLYIK 646

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3035	87.4	847	2 T09448	envelope glycoprot
2	3032	87.3	843	1 H44001	env polyprotein pr
3	3025	87.1	852	2 T12016	envelope glycoprot
4	3008	86.6	847	2 S13289	env protein - huma
5	2984	85.9	855	1 VCLJ42	env polyprotein pr
6	2953	85.0	861	1 VCLJ3C	env polyprotein pr
7	2939.5	84.6	856	1 VCLJH3	env polyprotein pr
8	2938.5	84.6	852	1 VCLJBR	env polyprotein -
9	2937.5	84.6	868	1 VCLJH4	env polyprotein -
10	2933	84.4	861	1 VCLJIV	env polyprotein pr
11	2928	84.3	851	2 S33985	env polyprotein -
12	2926.5	84.2	856	1 VCLJVL	env polyprotein pr
13	2914.5	83.9	854	2 S13288	env protein - huma
14	2908.5	83.7	729	1 VCLJXK	env polyprotein pr
15	2908.5	83.7	861	1 VCLJKE	env polyprotein pr
16	2894.5	83.3	856	1 VCLJ3W	env polyprotein pr
17	2890.5	83.2	859	1 VCLJMN	env polyprotein pr
18	2878.5	77.1	859	2 T01672	envelope polyprote
19	2672	76.9	855	1 VCLJZR	env polyprotein pr
20	2663	76.7	853	2 S54384	envelope polyprote
21	2632.5	75.8	856	1 A44963	env polyprotein pr
22	2611.5	75.2	846	1 VCLJND	env polyprotein pr
23	2196	63.2	854	1 VCLJSI	env polyprotein pr
24	2090.5	60.2	506	2 A40218	envelop glycoprote
25	1945.5	56.0	445	2 A41621	env polyprotein M
26	1941.5	55.9	443	2 C41621	env polyprotein P
27	1849	53.2	454	2 B41621	env polyprotein D
28	1848	53.2	495	2 S31493	env polyprotein -
29	1830.5	52.7	877	2 S49197	envelope protein p

ALIGNMENTS

RESULT 1

T09448

envelope glycoprotein - human immunodeficiency virus type 1 (strain JRPL)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999

C:Accession: T09448

R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,

submitted to the EMBL Data Library, July 1996

A:Reference number: Z16673

A:Accession: T09448

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-847 <PAN>

A:Cross-references: EMBL:U63632; NID:G1465777; PID:G1465781

C:Genetics:

C:Gene: env

C:Superfamily: type E retrovirus env polyprotein

Query Match	87.4%;	Score	3035;	DB 2;	Length	847;			
Best Local Similarity	85.4%;	Pred. No.	1.7e-211;						
Matches	577;	Conservative	29;	Mismatches	38;	Indels	32;	Gaps	6;
QY	1	MRVKGIRKYNQHLWRGGTLLGLMLTCSAVEKLVWTVYGVVWKEATTTLFCASDAKAY	60						
DB	1	MRVKGIRKSYQLWKGGTLLGLTLMICSAVEKLVWTVYGVVWKEATTTLFCASDAKAY	60						
QY	61	DTEVHVWATHACVPTDPNPQEIIVLENTVFNMWKNVQVQMHEDIISLWDQSLKPCVK	120						
DB	61	DTEVHVWATHACVPTDPNPQEVLENTVFHEFNWKNVQVQMHEDIISLWDQSLKPCVK	120						
QY	121	LTPLCVTLHCTNLKNATNTKSSNWKEMDRGEIKNCSFKVGAG-----	162						
DB	121	LTPLCVTLNCKDV-NATNTNDSEGMERGEIKNCSFNITTSIRDEVQKYEALFYKLDVV	179						
QY	163	-----KLINCNTSVITQACPKVSFPPIPIHYCAPAGFALLKCNKDFNKGSGPCTNV	213						
DB	180	PIDNNNTSYRLISCDTSVITQACPKISFPPIPIHYCAPAGFALLKCNKDFNKGSGPCKNV	239						
QY	214	STVQCTHGRIPVVSQTOLLNGSLAEEGVIRSENFTDNAKTIIVQLKESVEINCTRPNNN	273						
DB	240	STVQCTHGRIPVVSQTOLLNGSLAEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNN	299						
QY	274	TRKSIITIGRPFATYDGIIDIROAHCHNTSGEKWNTLKOIVTKLQAGFNKTIIVPKQS	333						
DB	300	TRKSHIGRPFATYTTGEIIGDIRQAHCHNISRAKWNDLTKQIVIKLREQENKTIIVFNHS	359						
QY	334	SGGDPPIVMHSFNCGGFFYCNSTQLFNSTW-NNTIGPNNTNG-TITLPCRICKQIINRWQ	391						
DB	360	SGGDPPIVMHSFNCGGFFYCNSTQLFNSTWNNNTGSGNNTEGNTITLPCRICKQIINMWQ	419						
QY	392	EYVKAMYPPIRQIRFCSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYK	451						

```

Db      420 EVCKAMVAPPPIRGQIRCSSNITGLLLTRDGGIN-ENGTEIFRPGGDMRDNRSELYKYK 478
Qy      452 VKIEPLGVAPTAKRRVVQREKRAVTLGAMFLGFLGAAGSTMGARSILITVQARQLLSG 511
Db      479 VKIEPLGVAPTAKRRVVQREKRAVIGAVFGLGFLGAAGSTMGAAASMTLTVQARQLLSG 538
Qy      512 IVQQNNLLRAIEAQOHLQLTWGIGK-LQARVLAVERYLKQOQLLGIWGCSGKLICTTA 570
Db      539 IVQQNNLLRAIEAQOQMLQLTWGIGK-LQARVLAVERYLGDQQLLGIWGCSGKLICTTA 598
Qy      571 VPMNASWSNKS LDOIWNNTMTMEWEREIDNYTNLYTLIEESQOQEKNEQELLELDKWA 630
Db      599 VPMNASWSNKS LDOIWNNTMTMEWEREIDNYTSEIVTLIEESQOQEKNEQELLELDKWA 658
Qy      631 SLWNWFDISKWLYYIK 646
Db      659 SLWNWFDITKWLWYIK 674

RESULT 2
H44001
env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994
C:Accession: H44001
R:Li, Y.; Hu, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.
J. Virol. 66, 6587-6600, 1992
A:Title: Complete nucleotide sequence, genome organization, and biological properties of
A:Reference number: A44001; MUID:93021387; PMID:1404605
A:Accession: H44001
A:Molecule type: DNA
A:Residues: 1-843 <LIY>
A:Cross-references: GB:M93258
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:19-35/Region: hydrophobic
F:30-489/Product: coat protein gp120 #status predicted <GP1>
F:490-843/Product: coat protein gp41 #status predicted <GP2>
F:499-515/Region: hydrophobic
F:673-689/Region: hydrophobic
F:738-755/Domain: transmembrane #status predicted <TMN>
F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Query Match      87.3%; Score 3032; DB 1; Length 843;
Best Local Similarity 84.2%; Pred. No. 2.8e-211;
Matches 567; Conservative 38; Mismatches 38; Indels 30; Gaps 5;

Qy      1 MRVKGIRKNYQHLWRGQTLGLLGLMICS AVEKLVVTVYVGVVWKEATTLFCASDAKAY 60
Db      1 MRATEIRKNYQHLWKGTLLGLLGLMICS AAEQLWTVYVGVVWKEATTLFCASDAKAY 60
Qy      61 DTEVHNWATHACVPTDPNPQETVLENVTENFNWKNMNVQEMHEDIISLDQSLKPCVK 120
Db      61 DTEVHNWATHACVPTDPNPQEVKLENVTENFNWKNMNVQEMHEDIISLDQSLKPCVK 120
Qy      121 LTPLCVTLHCTN-LKNATNTKSNWKEMDRGEIKNCSFKV----- 159
Db      121 LTPLCVTLNCTDLRNATNTSSSWETMEKGEIKNCSFNITTSRDVKQKEYALFYNLVV 180
Qy      160 ----GAGKLCNCTSVITQACPKVSEPIPIHYCAPAGFAILKCNCKKFGSGPCTNVST 215
Db      181 PIDNASYRLISCNSTSVITQACPKVSEPIPIHYCAPAGFAILKCNCKKFGTGPCTNVST 240
Qy      216 VQCTHGIRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRPNNTR 275
Db      241 VQCTHGIRPVVSTQLLNGSLABEEIVIRSENFDTNAKTIIVQLNESVVINCTRPNNTR 300
```

```

Qy      276 KSITIGPGRAFATGDIIGDIRQAHCNISGEKWNNTLKQIVTKLQAFG-NKTIVFKQSS 334
Db      301 KSNIGPGRALYTTGEIIGDIRQAHCNLSKTQWENTLEQIAIKLKEQFGNNKTIIFNPSS 360
Qy      335 GGDPEIVMHSFNCGGEFFYCNSTQLPNSTWNTIIGNNTNGITITLPCRITKQIINRQEVG 394
Db      361 GGDPEIVTSHSFNCGGEFFYCNSTQLF--TWNDRKLNNTGRNITLPCRITKQIINMWEVG 418
Qy      395 KAMYAPPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIPRPGGDMRDNRSELYKYKVKV 454
Db      419 KAMYAPPPIRGQIRCSSNITGLLLTRDGGKD-TNGTEIFRPGGDMRDNRSELYKYKVKV 477
Qy      455 IEPLGVAPTAKRRVVQREKRAVTLGAMFLGFLGAAGSTMGARSILITVQARQLLSGIVQ 514
Db      478 IEPLGVAPTAKRRVVQREKRAVGLGALFLGFLGAAGSTMGAAASITLTVQARQLLSGIVQ 537
Qy      515 QQNNLLRAIEAQOHLQLTWGIGK-LQARVLAVERYLKQOQLLGIWGCSGKLICTTAVPW 573
Db      538 QQNNLLRAIEAQOHLQLTWGIGK-LQARVLAVERYLKQOQLLGIWGCSGKLICTTVPW 597
Qy      574 NASWSNKS LDOIWNNTMTMEWEREIDNYTNLYTLIEESQOQEKNEQELLELDKWSLW 633
Db      598 NTSWSNKS LNEIWDNNTMTMEWEREIDNYTHIYSLIEESQOQEKNEQELLELDKWSLW 657
Qy      634 NWFDISKWLYYIK 646
Db      658 NWFDITKWLWYIK 670

RESULT 3
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12016
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S.
A:Reference number: Z17379; MUID:98178716; PMID:9519894
A:Accession: T12016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
A:Cross-references: EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g2351784
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match      87.1%; Score 3025.5; DB 2; Length 852;
Best Local Similarity 84.3%; Pred. No. 8.4e-211;
Matches 575; Conservative 34; Mismatches 34; Indels 39; Gaps 7;

Qy      1 MRVKGIRKNYQHLWRGQTLGLLGLMICS AVEKLVVTVYVGVVWKEATTLFCASDAKAY 60
Db      1 MRVKGIRKNQCHLWKWGTLLGLLGLMICS AAEQLRVTVYVGVVWKEATTLFCASDAKAY 60
Qy      61 DTEVHNWATHACVPTDPNPQETVLENVTENFNWKNMNVQEMHEDIISLDQSLKPCVK 120
Db      61 DTEVHNWATHACVPTDPNPQEVVLENVTENFNWKNMNVQEMHEDIISLDQSLKPCVK 120
Qy      121 LTPLCVTLHCTN-LKNATNTKSNWKEMDRGEIKNCSFKV----- 159
Db      121 LTPLCVTLNCTDYLRNDTNTSSNNGGMEGGEIKNCSFNITTRIGNKVKQKEYALFYKLDV 180
Qy      160 ----GAGKLCNCTSVITQACPKVSEPIPIHYCAPAGFAILKCNCKKFGSGPCTN 212
Db      181 VPIDNTTTSYRLINCNTSVITQACPKVSEPIPIHYCTPAGFALLKCKKFGTGPCTN 240
Qy      213 VSTVQCTHGIRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRPNN 272
Db      241 VSTVQCTHGIRPVVSTQLLNGSLABEEVVIRSENFDTNAKTIIVQLNESVEINCTRPNN 300
Qy      273 NTRKSTITIGPRAFYATGDIIGDIRQAHCNISGEKWNNTLKQIVTKLQAFQGNKTIIVFKQ 332
```

```
Db 301 NTRKSHIGPGQALYATGAIIGDIRQAHNCISRAKWNNTLKQIVKKLRVQFGNKTIIIFNQ 360
QY 333 SSGDPEIVMHSFNCGGEFFYCNSTOLFNSWTW--NNTIGNNNT----NGTITLPCRIRKQ 385
Db 361 SSGDPEIVMHSFNCGGEFFYCNSTLKFNSWTFNWTW--NDTEDEGNGTITLPCRIRKQ 418
QY 386 IINRWQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNWRS 445
Db 419 IINRWQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGTN--NSTNETFRPGGDMRDNWRS 477
QY 446 ELYKYKVKIEPLGVAPTAKRRVQREKRAVTLGAMFGLGFLGAAGSTWGARSLLITVQA 505
Db 478 ELYKYKVKIEPLGVAPTAKRRVQREKRAVTLGAMFGLGFLGAAGSTWGAAAVTLITVQA 537
QY 506 RQLLSGIVQOONNLLRAIEAQHLLQLTVWGIK--LQARVLAVERYLKDOQLLGIWCSGK 564
Db 538 RQLLSGIVQOONNLLRAIEAQHLLQLTVWGIK--LQARVLAVERYLRDQQLLGIWCSGK 597
QY 565 LICTTAVPNWASNSKSLDQIMNNMTWMEWERIDNYTLIYTLIESQOQEKNEQELL 624
Db 598 LICTTAVPNWASNSKSLDQIMNNMTWMEWERIDNYTLIYTLIESQOQVQNEOELL 657
QY 625 ELDKWSLWNSWFDISKWLWYIK 646
Db 658 ELDKWSLWNSWFDITKWLWYIK 679

RESULT 4
S13289
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S13289
R.O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A
Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13289
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <OR>
C:Superfamily: type E retrovirus env polyprotein

Query Match 86.6%; Score 3008; DB 2; Length 847;
Best Local Similarity 84.6%; Pred. No. 1.5e-209;
Matches 572; Conservative 29; Mismatches 43; Indels 32; Gaps 6;

QY 1 MRVKGIRKQYQHLWRGGTLLGLMLICSAVEKLVWTVVYGVVPWKBATTLTFCASDAKAY 60
Db 1 MRVKGIRKSYQLWKWGTMLLGLMLICSAVEKLVWTVVYGVVPWKBATTLTFCASDAKAY 60
QY 61 DTEVHNWATHACVPTDNPQEIIVLNTENFNWKNMKNVQEHEDIISLDQSLKPCVK 120
Db 61 DTEVHNWATHACVPTDNPQEIIVLNTENFNWKNMKNVQEHEDIISLDQSLKPCVK 120
QY 121 LTPLCVTLHCTNLKATNTKSSNWKEMDRGEIKNCSFKVKGAG----- 162
Db 121 LTPLCVTLNCKDV--NATNTNDSGTMERGEIKNCSFNITTSIRDEQVEYALFYKLDVV 179
QY 163 -----KLINCNTSVITQACPKVSFPIPIHYCAPAGFAILKCNCKKFNKSGPCTNV 213
Db 180 PIDNNNTSVRLISCDTSVITQACPKISFPIPIHYCAPAGFAILKCNCKTENGKGPCKNV 239
QY 214 STVQCTHGIRPVVSTOLLNGLSABEGVIRSENFTDNAKTIIIVQLKESVEINCTPNNN 273
Db 240 STVDCTHGIRPVVSTOLLNGLSABEEVIRSDNFTNNAKTIIVQLKESVEINCTPNNN 299
QY 274 TRKSIITIGPRAFYATGDIIGDIRQAHNCISGKMNNTLKQIVTKLQAOFGNKTIVFKOS 333
Db 300 TRKSHIHGPRAFYTTGDIIGDIRQAHNCISRAKWNNTLKQIVIKURBQFNKTIIVFNS 359
QY 334 SSGDPEIVMHSFNCGGEFFYCNSTOLFNSWTW--NNTIGNNNTNG--TITLPCRIRKQIINRWQ 391
```

```
Db 360 SSGDPEIVMHSFNCGGEFFYCNSTOLFNSWTWNNTEGSSNNTTEGNTITLPCRIRKQIINRWQ 419
QY 392 EYKAMYPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNWRSLEYKYK 451
Db 420 EYKAMYPPIRGQIRCSSNITGLLLTRDGGIN--ENGTEIFRPGGDMRDNWRSLEYKYK 478
QY 452 VVKIEPLGVAPTAKRRVQREKRAVTLGAMFGLGFLGAAGSTMGARSLLITVQARQLLSG 511
Db 479 VVKIEPLGVAPTAKRRVQREKRAVIGAGVFLGFLGAAGSTMGAAASMTLTVQARQLLSG 538
QY 512 IVQOQNLLRAIEAQHLLQLTVWGIK--LQARVLAVERYLKDOQLLGIWCSGKLCITTA 570
Db 539 IVQOQNLLRAIEAQOQMLQLTVWGIK--LQARVLAVERYLQDQLLGIWCSGKLCITTA 598
QY 571 VPMNASWSNSKSLDQIMNNMTWMEWERIDNYTLIYTLIESQOQEKNEQELLEDKWA 630
Db 599 VPMNASWSNSKSLDQIMNNMTWMEWERIDNYTLIYTLIESQOQEKNEQELLEDKWA 658
QY 631 SLWNMFDISKWLWYIK 646
Db 659 SLWNMFDISKWLWYIK 674

RESULT 5
VCLJAZ
env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03976
R.Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-Shi
Science 227, 484-492, 1985
A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).
A:Reference number: A04003; MUID:85090453; PMID:2578227
A:Accession: A03976
A:Molecule type: DNA
A:Residues: 1-855 <SAN>
A:Cross-references: GB:K02007; NID:g328658; PIDN:AAB59882.1; PID:g328666
C:Genetics:
C:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprote
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-509/Product: exterior membrane glycoprotein #status predicted <EXT>
F:510-855/Product: transmembrane glycoprotein #status predicted <TM>
F:87,129,140,158,184,190,200,244,265,292,298,304,334,341,358,364,388,394,400,408,445,458,
F:610,624,636,815/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.9%; Score 2984; DB 1; Length 855;
Best Local Similarity 82.4%; Pred. No. 8.5e-208;
Matches 563; Conservative 34; Mismatches 48; Indels 38; Gaps 6;

QY 1 MRVKGIRKQYQHLWRGGTLLGLMLICSAVEKLVWTVVYGVVPWKBATTLTFCASDAKAY 60
Db 1 MKVKGTRRNYQHLWRWGTLLGLMLICSAVEKLVWTVVYGVVPWKBATTLTFCASDAKAY 60
QY 61 DTEVHNWATHACVPTDNPQEIIVLNTENFNWKNMKNVQEHEDIISLDQSLKPCVK 120
Db 61 DTEVHNWATHACVPTDNPQEIIVLNTENFNWKNMKNVQEHEDIISLDQSLKPCVK 120
QY 121 LTPLCVTLHCTNLKATNTKSSNWKEMDRGEIKNCSFKVKGAG----- 162
Db 121 LTPLCVTLNCTDLGRATNTNNSNWKKEIKGEIKNCSFNITTSIRDKIQENALFRMLDVV 180
QY 163 -----KLINCNTSVITQACPKVSFPIPIHYCAPAGFAILKCNCKKFNKSG 208
Db 181 PIDNASTTNTYNYRLIHCNRSVITQACPKVSFPIPIHYCAPAGFAILKCNCKTENGK 240
QY 209 PCTNVSTVQCTHGIRPVVSTOLLNGLSABEGVIRSENFTDNAKTIIIVQLKESVEINCT 268
Db 241 PCTNVSTVQCTHGIRPVVSTOLLNGLSABEEVIRSDNFTNNAKTIIVQLNESVAINCT 300
```

```

QY 269 RPNNTKSTITGPGAFYATGDIIGDIRQAHNCISGEKNNNTLKOIVTKLOAQFG-NKT 327
DB 301 RPNNTKSTIYIGPGAFHTTGRIGDIRKHAHCNISRAQWNTTLEQIVKKLREQFGNNKT 360
QY 328 IVFKQSSGGDPETVMSHFNCGGFFCYCNSTQLFNSTW--NNTIGPNNNTNGTITLPCRIKQ 385
DB 361 IVFNQSSGGDPETVMSHFNCGGFFCYCNSTQLFNSTWLNHTGK-NDTIILPCRIKQ 419
QY 386 IINRQEVGKAMYAPPIRGQIRCSSNITGLLTRDGGKEISNTTEIFRPGGMDRDNWRS 445
DB 420 IINRQEVGKAMYAPPIRGQIRCSSNITGLLTRDGGKEISNTTEIFRPGGMDRDNWRS 479
QY 446 ELYKVKVIEPLGVAPTAKRVVQREKRAV-TLGAMFLGLFAGAGSTMGARSLLTVQ 504
DB 480 ELYKVKVIEPLGVAPTAKRVVQREKRAV-TLGAMFLGLFAGAGSTMGARSLLTVQ 539
QY 505 ARQLLSGIVQOQNNLLRAEAQHLLQLTWVGIK-LQARVLAVERYLKDQQLLGIWCSG 563
DB 540 ARQLLSGIVQOQNNLLRAEAQHLLQLTWVGIK-LQARVLAVERYLKDQQLLGIWCSG 599
QY 564 KLICTTAVPNASWSNKSLDQIWNNTMWEWEREIDNYTNLIYTLLEESQOQKNEQEL 623
DB 600 KLICTTAVPNASWSNKSLEIDWNTMWEWEREIDNYTNLIYTLLEESQOQKNEQEL 659
QY 624 LELDKWASLWNPDISKWLWYIK 646
DB 660 LELDKWASLWNPDISKWLWYIK 682

RESULT 6
VCLJSC
env polypotein precursor - human immunodeficiency virus type 1 (isolate SC)
N:Alternate names: coat polypotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Accession: B28922
F:Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Staal, A.;
Virology 164, 531-536, 1988
A:Title: Envelope sequences of two new United States HIV-1 isolates.
A:Reference number: A28922; MUID:88219542; PMID:3369091
A:Accession: B28922
A:Molecule type: DNA
A:Residues: 1-861 <GUR>
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypotein
C:Keywords: capsid protein; coat protein; glycoprotein; polypotein; transmembrane prote
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-861/Product: env polypotein #status predicted <EP>
F:87,129,135,140,143,159,163,187,198,234,241,262,276,295,301,302,333,340,356,362,386,396

Query Match 85.0%; Score 2953; DB 1; Length 861;
Best Local Similarity 82.1%; Pred. No. 1.5e-205;
Matches 566; Conservative 33; Mismatches 46; Indels 44; Gaps 9;

QY 1 MRVKGIRKNOYHLMRGGTLGLMLMICSAAVEKLVTVVYGVVPVKWKAATTLFPCASDAKAY 60
DB 1 MRVKGIRKNOYHLMRGGTLGLMLMICSAAVEKLVTVVYGVVPVKWKAATTLFPCASDAKAY 60
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENTNFENNMKNVQEHEDIISLWDSQKPCVK 120
DB 61 DTEVHNWATHACVPTDPNPQEIIVLENTNFENNMKNVQEHEDIISLWDSQKPCVK 120
QY 121 LTPCLVTLHCTNLK-----NATNTKSNWKEMDRGEIKNCSFKVAGK----- 163
DB 121 LTPCLVTLHCTNLK-----NATNTKSNWKEMDRGEIKNCSFKVAGK----- 163
QY 164 -----LINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKFKNSGSPC 210
DB 181 DVVPIDNDTTSYTLINCNSTVITQACPKVSFEPIPIHYCA-RWFAILCNCKFKNSGTPC 239

```

```

QY 211 TNVSTVQCTHGIRPVVSTQLLNGSLABEGVVIRSENFSTDAKTIIVQLKESVEINCTRP 270
DB 240 TNVSTVQCTHGIRPVVSTHLLNGSLABEEVVLRSNFSTDAKTIIVQLKEAVEINCTRP 299
QY 271 NNNTKSTITI--GPGRAFAYATGDIIGDIRQAHNCISGEKNNNTLKOIVTKLOAQFGNKT 328
DB 300 NNNTTRSIIHQRPGRFAYATGDIIGDIRQAHNCISRAKNNNTLKOIVTKLRDQENKTI 359
QY 329 VFKQSSGGDPETVMSHFNCGGFFCYCNSTQLFNSTWNTTIGP--NNTNG--TITLPCRI 383
DB 360 IFNRSSGGDPETVMSHFNCGGFFCYCNSTQLFNSTWNTTIGP--NNTNG--TITLPCRI 419
QY 384 KOIINRQEVGKAMYAPPIRGQIRCSSNITGLLTRDGGKEISNT-----TEIFRPGGMD 439
DB 420 KEIINRQEVGKAMYAPPIRGQIRCSSNITGLLTRDGGKSNKSNKSNKSNKSNKSNKSNK 479
QY 440 RDNWSELYKVKVIEPLGVAPTAKRVVQREKRAV-TLGAMFLGLFAGAGSTMGARS 498
DB 480 RDNWSELYKVKVIEPLGVAPTAKRVVQREKRAV-TLGAMFLGLFAGAGSTMGARS 539
QY 499 LTLTVQARQLLSGIVQOQNNLLRAEAQHLLQLTWVGIK-LQARVLAVERYLKDQQLLG 557
DB 540 MTLTVQARQLLSGIVQOQNNLLRAEAQHLLQLTWVGIK-LQARVLAVERYLKDQQLLG 599
QY 558 IWGCSGKLICTTAVPNASWSNKSLDQIWNNTMWEWEREIDNYTNLIYTLLEESQOQE 617
DB 600 IWGCSGKLICTTAVPNASWSNKSLDQIWNNTMWEWEREIDNYTNLIYTLLEESQOQE 659
QY 618 KNEQELLELDKWLWNPDISKWLWYIK 646
DB 660 KNEQELLELDKWLWNPDISKWLWYIK 688

RESULT 7
VCLJSH
env polypotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)
N:Alternate names: coat polypotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Accession: A03973
F:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Doraz
nberger, J.A.; Papas, T.S.; Ghrayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.
Nature 313, 277-284, 1985
A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
A:Reference number: A93353; MUID:85111123; PMID:2578615
A:Accession: A03973
A:Molecule type: DNA
A:Residues: 1-856 <RAT>
A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:G326383; PIDN:AAA4442(
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypotei
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406,
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predict
Query Match 84.6%; Score 2939.5; DB 1; Length 856;
Best Local Similarity 82.6%; Pred. No. 1.4e-204;
Matches 568; Conservative 34; Mismatches 39; Indels 47; Gaps 12;

QY 1 MRVKGIRKNOYHLMRGGTLGLMLMICSAAVEKLVTVVYGVVPVKWKAATTLFPCASD 56
DB 1 MRVKGIRKNOYHLMRGGTLGLMLMICSAAVEKLVTVVYGVVPVKWKAATTLFPCASD 57
QY 57 AKAYDEVHNWATHACVPTDPNPQEIIVLENTNFENNMKNVQEHEDIISLWDSQK 116
DB 58 AKAYDEVHNWATHACVPTDPNPQEIIVLENTNFENNMKNVQEHEDIISLWDSQK 117
QY 117 PCVKLTPLCVTLHCTNLKSNATNTKSNWKE-MDRGEIKNCSFKVGA---GK----- 163

```

Db 118 PCVKLTPLCVSLKCTDKNDTNTSSGRIKMGKIKNSFNISIRKVOKEYAFFY 177  
Qy 164 -----LINCNTSVITQACPKVSEPIPIHYCAPAGFAILLKNDKKNFGSG 208  
Db 178 KLDIIPIDNDTSTYLTSCNTSVITQACPKVSEPIPIHYCAPAGFAILLKKNKTFNGTG 237  
Qy 209 PCTNVSTVQCTHGRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVOLKESVEINCT 268  
Db 238 PCTNVSTVQCTHGRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVOLNQSVEINCT 297  
Qy 269 RPNNTKXSIIT--GRCRAFAVATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLOAOPG-N 325  
Db 298 RPNNTKXSIIRIQGPGRAFTVIGK-IGNMRQAHCHNISRAKWNNTLKQIDSKLREOFGN 356  
Qy 326 KTIIVFKQSSGDDPEIVMHSFNCGGEFFYCNSTOLFSTWNN---TIGPNNTNG--TITL 379  
Db 357 KTIIVFKQSSGDDPEIVMHSFNCGGEFFYCNSTOLFSTWNTSTWSTKSGNNTSGSDTITL 416  
Qy 380 PCRIKQIINQVGVKAMYPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGDM 439  
Db 417 PCRIKQIINQVGVKAMYPPIRGQIRCSNITGLLLTRDGGNS--NNESEIFRPGGDM 475  
Qy 440 RDNWRSELYKVKVKEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAGSTMGARS 499  
Db 476 RDNWRSELYKVKVKEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAGSTMGAAS 535  
Qy 500 TLTVQARQLSGIVQOONLLRAIEAQHLLQTLTVNGIK--LQARVLAVERYLKDQQLLGI 558  
Db 536 TLTVQARQLSGIVQOONLLRAIEAQHLLQTLTVNGIKQOLARILAVERYLKDQQLLGI 595  
Qy 559 WCGSGKLICTTAVPWNASNSKSLDQIWNMTWMEWEREIDNTNLIYTLIEBSQNOQEK 618  
Db 596 WCGSGKLICTTAVPWNASNSKSLDQIWNMTWMEWEREIDNTNLIYTLIEBSQNOQEK 655  
Qy 619 NEQELLELDKWSLWVDFISKWLWYIK 646  
Db 656 NEQELLELDKWSLWVDFISKWLWYIK 683

## RESULT 8

VCLJH4

env polyprotein - human immunodeficiency virus type 1 (isolate BR)

N;Alternate names: coat polyprotein

N;Contains: coat protein gp120; coat protein gp41

C;Species: human immunodeficiency virus type 1, HIV-1

A;Note: host Homo sapiens (man)

C;Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Feb-1997

C;Accession: A31667

R;Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S.

Virology 168, 79-89, 1989

A;Title: Biological and molecular characterization of human immunodeficiency virus (HIV-

A;Reference number: A94389; MUID:89085613; PMID:2789516

A;Accession: A31667

A;Molecule type: DNA

A;Residues: 1-852 &lt;ANA&gt;

C;Superfamily: type E retrovirus env polyprotein

C;Keywords: capsid protein; coat protein; polyprotein; transmembrane protein

F;1-516/Product: coat protein gp120 #status predicted &lt;CP1&gt;

F;517-852/Product: coat protein gp41 #status predicted &lt;CP2&gt;

Query Match 84.6%; Score 2938.5; DB 1; Length 852;

Best Local Similarity 81.6%; Pred. No. 1.7e-204;

Matches 556; Conservative 41; Mismatches 47; Indels 37; Gaps 8;

Qy 1 MRVKGIRKNTYQHLWR--GGTLLGLMLICSAVEKLWTVVYGVVWKEATTLFCASDAKA 59

Db 1 MRVKGIRKNTYQHLWR--GGTLLGLMLICSAVEKLWTVVYGVVWKEATTLFCASDAKA 60

Qy 60 YDTEHNWVATHACVTPDNPQELVLENVTENFMKKNVVEOMHEDIISLWDSQSKPCV 119

Db 61 YDTEHNWVATHACVTPDNPQELVLENVTENFMKKNVVEOMHEDIISLWDSQSKPCV 120

Qy 120 KLTPLCVTLHCTNLKNAITKSNKEMDRGGEIKKNSCFKVGAG----- 162  
Db 121 KLTPLCVTLHCTNLKNAITKSNKEMDRGGEIKKNSCFKVGAG----- 179  
Qy 163 -----KLINCNTSVITQACPKVSEPIPIHYCAPAGFAILLKNDKKNFGSGPC 210  
Db 180 VPIDNDKNTWTRKRLISCNSTSVITQACPKVSEPIPIHYCAPAGFAILLKKNKTFNGTGPC 239  
Qy 211 TNVSTVQCTHGRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVOLKESVEINCTRP 270  
Db 240 TNVSTVQCTHGRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVOLNQSVEINCTRP 299  
Qy 271 NNTRKSIITIGPGRAPYATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLOAOPGNTIVF 330  
Db 300 NNTRKSIITIGPGRAPYATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLOAOPGNTIVF 359  
Qy 331 KQSSGDDPEIVMHSFNCGGEFFYCNSTOLFSTW--NNTIGPNNTNG--ITLPCRKQI 387  
Db 360 NRSQSGDDPEIVMHSFNCGGEFFYCNSTOLFSTWYNTTG--NITEGNSPITLPCRKQI 418  
Qy 388 NRQEVGKAMYPPIRGQIRCSNITGLLLTRDGGKEISNT--TEIFRPGGDMRDNWRSE 446  
Db 419 NRQEVGKAMYPPIRGQIRCSNITGLLLTRDGGNNNETTDEIFRPGGDMRDNWRSE 478  
Qy 447 LYKYKVVKEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAGSTMGARSITLTVQAR 506  
Db 479 LYKYKVVKEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAGSTMGARSITLTVQAR 538  
Qy 507 QLLSGIVQOONLLRAIEAQHLLQTLTVNGIK--LQARVLAVERYLKDQQLLGIWCGSGKL 565  
Db 539 LLLSGIVQOONLLRAIEAQHLLQTLTVNGIKQOLARVLAVERYLKDQQLLGIWCGSGKL 598  
Qy 566 ICTTAVPWNASNSKSLDQIWNMTWMEWEREIDNTNLIYTLIEBSQNOQEKNEQELLE 625  
Db 599 ICTTAVPWNASNSKSLDQIWNMTWMEWEREIDNTNLIYTLIEBSQNOQEKNEQELLE 658  
Qy 626 LDKWASLWVDFISKWLWYIK 646  
Db 659 LDKWASLWVDFISKWLWYIK 679

## RESULT 9

VCLJH4

env polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)

N;Alternate names: coat polyprotein

N;Contains: coat protein gp120; coat protein gp41

C;Species: human immunodeficiency virus type 1, HIV-1

C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 16-Jul-1999

C;Accession: C25523

R;Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devare, S.

Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986

A;Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human

A;Reference number: A94136; MUID:87041461; PMID:3490666

A;Accession: C25523

A;Molecule type: DNA

A;Residues: 1-868 &lt;DES&gt;

A;Cross-references: GB:M13137; NID:G326460; PIDN:AAA44311.1; PID:G326467

C;Genetics:

A;Gene: env

C;Superfamily: type E retrovirus env polyprotein

C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei

F;1-521/Product: coat protein gp120 #status predicted &lt;GP1&gt;

F;522-868/Product: coat protein gp41 #status predicted &lt;GP2&gt;

F;89,138,139,142,166,195,198,208,245,252,273,287,300,312,342,349,365,371,395,405,409,459,

Query Match 84.6%; Score 2937.5; DB 1; Length 868;

Best Local Similarity 79.5%; Pred. No. 2e-204;

Matches 556; Conservative 37; Mismatches 47; Indels 59; Gaps 8;

Qy 1 MRVKGIRKNTYQHLWR--GGTLLGLMLICSAVEKLWTVVYGVVWKEATTLFCASDAKA 60

Db 3 MRVKGIRKNTYQHLWR--GGTLLGLMLICSAVEKLWTVVYGVVWKEATTLFCASDAKA 62

```
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENTNFNMKNMKNVQEMHEDIISLWQSLKPCVK 120
Db 63 DTEAHNVWATHACVPTNPQEVVLENTNFNMKNMKNVQEMHEDIISLWQSLKPCVK 122
QY 121 LTPCLVTLHCT--NLKNATNKSNN-----WKMDRGEIKNCSFKVAG-----162
Db 123 LTPCLVTLNCTDLNTNTTTELSIIVVMEQRGKGEMRNCNFIITTSIRDVKQREYALF 182
QY 163 -----KLINCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNCK 202
Db 183 YKLDVPEIDDKNTNTNTKLLINCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNCK 242
QY 203 KFGSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVQLKES 262
Db 243 KFGNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVQLNVS 302
QY 263 VEINCTRPNNNTKRSITIGRPAFYATGDIIGDIRQAHGNISGEKWNNTLKQIVTKLQAO 322
Db 303 VEINCTRPNNHTKRVTLGPGRVYTTGELIGNIROAHCNISRAQWNNTLQOIATTLREQ 362
QY 323 FGNKTIIVFKOSSGGDPEIVMHSFNCGGEFFYCNSTOLFNSWNTTIGPNNNTGT-----376
Db 363 FGNKTIIVFKOSSGGDPEIVMHSFNCGGEFFYCNSTOLFNSAMNVT-----SNGTWSVTRK 417
QY 377 -----ITLPCRIKQIINRWQSVGKAMYAPPIRGQIRCSSNITGLLLTRDGGKEISNTT 429
Db 418 OKDTGDIITLPCRIKQIINRWQSVGKAMYALPIKGLIRCSSNITGLLLTRDGGGE-NQIT 476
QY 430 EIPRPGGDMRDNRSELYKYKVKIEPLGVAPTKAKRRVVOREKRAV-TLGMAMFLGFLG 488
Db 477 EIPRPGGDMRDNRSELYKYKVKIEPLGVAPTKAKRRVVOREKRAVGLGMAMFLGFLG 536
QY 489 AAGSTWGARSLLITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIK-LOARVLAVE 547
Db 537 AAGSTWGARSLLITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIKLOARILAVE 596
QY 548 RYLDQOQLLGWCSGKLICTTAVPNWASWSKSLDQIWNNTMTMEWERIDNNTLYIT 607
Db 597 RYLDQOQLLGWCSGKLICTTAVPNWASWSKSLDQIWNNTMTMEWERIDNNTLYIT 656
QY 608 LIBESQOQKNEQELLELDKWSLWNPFDISKWLWYIK 646
Db 657 LIBESQOQKNEQELLELDKWSLWNPFDISKWLWYIK 695

RESULT 10
VCLJLV
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03975
R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.
Cell 40, 9-17, 1985
A:Title: Nucleotide sequence of the AIDS virus, LAV.
A:Reference number: A90866; MUID:85099333; PMID:2981635
A:Accession: A03975
A:Molecule type: DNA
A:Residues: 1-861 <WAI>
A:Cross-references: GB:K02013; NID:g326417; PIDN:AAB59751.1; PID:g326424
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-516/Product: exterior membrane glycoprotein #status predicted <EXT>
F;517-861/Product: transmembrane glycoprotein #status predicted <TM>
F;88,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411
F;616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.4%; Score 2933; DB 1; Length 861;
Best Local Similarity 81.8%; Pred. No. 4.2e-204;
Matches 564; Conservative 34; Mismatches 42; Indels 44; Gaps 11;
```

```
Matches 567; Conservative 33; Mismatches 41; Indels 52; Gaps 12;
QY 1 MRVKGIRKKNYQHLWR-----GGTLLGLMLMICSAGEKLVWTVVYGVVPWKEATITLFCASD 56
Db 1 MRVK---EKYQHLRWGKNGTLLGLMLMICSATEKLVWTVVYGVVPWKEATITLFCASD 57
QY 57 AKAYDEVHNWATHACVPTDPNPQEIIVLENTNFNMKNMKNVQEMHEDIISLWQSLK 116
Db 58 AKAYDEVHNWATHACVPTDPNPQEVVLENTNFNMKNMKNVQEMHEDIISLWQSLK 117
QY 117 PCVKLTPLCVTLHCTNLKNATNTKSSNWE-----MDRGEIKNCSFKVGA---GK---163
Db 118 PCVKLTPLCVSLKCTDLGNATNTSNTSSGEMMMEKGEIKNCSFNITSIRGKVQKE 177
QY 164 -----KLINCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNCK 203
Db 178 YAFFYKLDIIPIDNDTTSYTLSCNTSVITQACPKVSPFPIPIHYCAPAGFAILKCNCK 237
QY 204 FNGSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVQLKESV 263
Db 238 FNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEGVVIRSENFDTNAKTIIVQLNQS 297
QY 264 EINCTRPNNNTKRSITI--GPGRAFATGDIIGDIRQAHGNISGEKWNNTLKQIVTKLQ 321
Db 298 EINCTRPNNNTKRSIRIQRGPRAFVTIGK-IGNMRQAHGNTSRAKWNATLKQIASKLRE 356
QY 322 QFG-NKTIIVFKOSSGGDPEIVMHSFNCGGEFFYCNSTOLFNSWNTN---TIGPNNNTG- 375
Db 357 QFGNNKTIIVFKOSSGGDPEIVTHSFNCGGEFFYCNSTOLFNSWNTSSTWSTESNNTG 416
QY 376 -TITLPCRIKQIINRWQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGKEISNTTIFRP 434
Db 417 DTITLPCRIKQFINMWQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGNN-NNGSBIFRP 475
QY 435 GGGDMRDNRSELYKYKVKIEPLGVAPTKAKRRVVOREKRAVTLGAMFLGFLGAAGSTM 494
Db 476 GGGDMRDNRSELYKYKVKIEPLGVAPTKAKRRVVOREKRAVIGALFLGFLGAAGSTM 535
QY 495 GARSLLITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIK-LOARVLAVERYLKQ 553
Db 536 GARSMTLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIKLOARILAVERYLKQ 595
QY 554 QLLGIWCSGKLICTTAVPNWASWSKSLDQIWNNTMTMEWERIDNNTLYITLIESQ 613
Db 596 QLLGIWCSGKLICTTAVPNWASWSKSLDQIWNNTMTMEWERIDNNTLYITLIESQ 655
QY 614 NOEKNEQELLELDKWSLWNPFDISKWLWYIK 646
Db 656 NOEKNEQELLELDKWSLWNPFDISKWLWYIK 688

RESULT 11
S33985
env polyprotein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999
C:Accession: S33985
R:Carlini, F.
submitted to the EMBL Data Library, November 1991
A:Reference number: S33979
A:Accession: S33985
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-851 <CAR>
A:Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77628.1; PID:g60199
C:Superfamily: type E retrovirus env polyprotein

Query Match 84.3%; Score 2928; DB 2; Length 851;
Best Local Similarity 82.5%; Pred. No. 9.5e-204;
Matches 564; Conservative 34; Mismatches 42; Indels 44; Gaps 11;
```



Db 1 MRVK---EKYQHLWRGWGTMLLGMLMICSATEKLVTVVYGVVPWKAEATTLFCASD 57  
QY 57 AKAYDEVHNVWATHACVPTDPNPQEIIVLENVTENFMNMKNVVEQMHEDIISLWQSLK 116  
Db 58 AKAYDEVHNVWATHACVPTDPNPQEIIVLENVTENFMNMKNVVEQMHEDIISLWQSLK 117  
QY 117 PCVKLTPLCVTLHCTMKNATNTKSSNWE-MDRGEIKKCSFKVGAGK----- 163  
Db 118 PCVKLTPLCVSLKCTDLKNDTNTSSSGRMIMEKEIKKCSFNISTSKRGKVOKEYAFFY 177  
QY 164 -----LINCNTSVITQACPKVSEPIPIHYCAPAGPAIILKCNKDNKNGSG 208  
Db 178 KLDIIPIDNDTTSYTLTSCNTSVITQACPKVSEPIPIHYCAPAGPAIILKCNKDNKNGSG 237  
QY 209 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEVEGVIRSENFDTNAKTIIVQLKESVEINCT 268  
Db 238 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEVEGVIRSENFDTNAKTIIVQLKESVEINCT 297  
QY 269 RPNNTTRKSIITI--GGRAFYATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLOAQFG-N 325  
Db 298 RPNNTTRKKIRIQRGPGRAFVTIGK-IGNMRQAHCHNISRAKWNNTLKQIDSKLREQFGNN 356  
QY 326 KTIIVFKQSSGGDEPEIVMHSFNCGGEFFYCNSTOLFNSNTWNTTIGPNNNTG--TITLPCR 383  
Db 357 KTIIFKQSSGGDEPEIVTHSFNCGGEFFYCNSTOLFNSNTW-STKGSNNTEGSDTITLPCR 415  
QY 384 KQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRPGGGMDRDNW 443  
Db 416 KQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRDGGNS--NNESEIFRPGGGMDRDNW 474  
QY 444 RSLEYKYKVKIEPLGVAPTKAKRRVVOREKRAVTLGMPFLGAGSTMGARSITLTV 503  
Db 475 RSLEYKYKVKIEPLGVAPTKAKRRVVOREKRAVTLGMPFLGAGSTMGARSITLTV 534  
QY 504 QARQLLSGIVQOQNLLRAIEAQOHLQLTVWGK-LOARVLAVERYLKDQQLLGWGS 562  
Db 535 QARQLLSGIVQOQNLLRAIEAQOHLQLTVWGIKQLQARILAVERYLKDQQLLGWGS 594  
QY 563 GKLICTTAVPNWASWSNKSIDQIWNMTWMEWEREIDNTNLIYTLIESQOQKEQE 622  
Db 595 GKPICTTAVPNWASWSNKSLEQIWNMTWMEWEREIDNTNLIYTLIESQOQKEQE 654  
QY 623 LLELDKWSLWVFNFDISKWLWYIK 646  
Db 655 LLELDKWSLWVFNFDISKWLWYIK 678  
RESULT 12  
VCLJVL  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)  
N;Alternate names: coat polyprotein  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
R;Accession: A03974  
R;Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.  
Nature 313, 450-458, 1985  
A;Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retroviral genome  
A;Reference number: A93355; MUID:85111157; PMID:2982104  
A;Accession: A03974  
A;Molecule type: DNA  
A;Residues: 1-856 <MUE>  
A;Cross-references: GB:K02083; NID:g555008; PIDN:AAB59873.1; PID:g328559  
C;Gene: env  
C;Superfamily: type B retrovirus env polyprotein  
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein  
F;1-30/Domain: signal sequence #status predicted <SIG>  
F;31-511/Product: exterior membrane glycoprotein #status predicted <EXT>  
F;512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F;88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406  
F;611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.2%; Score 2926.5; DB 1; Length 856;  
Best Local Similarity 82.1%; Pred. No. 1.2e-203;  
Matches 565; Conservative 35; Mismatches 41; Indels 47; Gaps 12;  
QY 1 MRVKGIKKNYQHL---WRGGTLLGLMMLCSAVEKLVTVVYGVVPWKAEATTLFCASD 56  
Db 1 MRVK---EKYQHLWRGWGTMLLGMLMICSATEKLVTVVYGVVPWKAEATTLFCASD 57  
QY 57 AKAYDEVHNVWATHACVPTDPNPQEIIVLENVTENFMNMKNVVEQMHEDIISLWQSLK 116  
Db 58 AKAYDEVHNVWATHACVPTDPNPQEIIVLENVTENFMNMKNVVEQMHEDIISLWQSLK 117  
QY 117 PCVKLTPLCVTLHCTMKNATNTKSSNWE-MDRGEIKKCSFKVGA-----GK----- 163  
Db 118 PCVKLTPLCVSLKCTDLKNDTNTSSSGRMIMEKEIKKCSFNISTSKRGKVOKEYAFFY 177  
QY 164 -----LINCNTSVITQACPKVSEPIPIHYCAPAGPAIILKCNKDNKNGSG 208  
Db 178 KLDIIPIDNDTTSYTLTSCNTSVITQACPKVSEPIPIHYCAPAGPAIILKCNKDNKNGSG 237  
QY 209 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEVEGVIRSENFDTNAKTIIVQLKESVEINCT 268  
Db 238 PCTNVSTVQCTHGIRPVVSTQLLNGSLAEVEGVIRSENFDTNAKTIIVQLKESVEINCT 297  
QY 269 RPNNTTRKSIITI--GGRAFYATGDIIGDIRQAHCHNISGEKWNNTLKQIVTKLOAQFG-N 325  
Db 298 RPNNTTRKKIRIQRGPGRAFVTIGK-IGNMRQAHCHNISRAKWNNTLKQIDSKLREQFGNN 356  
QY 326 KTIIVFKQSSGGDEPEIVMHSFNCGGEFFYCNSTOLFNSNTWNTTIGPNNNTG--TITLPCR 379  
Db 357 KTIIFKQSSGGDEPEIVTHSFNCGGEFFYCNSTOLFNSNTWNTTIGPNNNTG--TITLPCR 416  
QY 380 PCRIKQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRPGGGDM 439  
Db 417 PCRIKQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRDGGNN--NNESEIFRPGGGDM 475  
QY 440 RDNWRSELYKYKVKIEPLGVAPTKAKRRVVOREKRAVTLGMPFLGAGSTMGARSIL 499  
Db 476 RDNWRSELYKYKVKIEPLGVAPTKAKRRVVOREKRAVTLGMPFLGAGSTMGARSIL 535  
QY 500 TLTVQARQLLSGIVQOQNLLRAIEAQOHLQLTVWGK-LOARVLAVERYLKDQQLLGI 558  
Db 536 TLTVQARQLLSGIVQOQNLLRAIEAQOHLQLTVWGIKQLQARILAVERYLKDQQLLGI 595  
QY 559 WGCSEGLICTTAVPNWASWSNKSIDQIWNMTWMEWEREIDNTNLIYTLIESQOQKEK 618  
Db 596 WGCSEGLICTTAVPNWASWSNKSLEQIWNMTWMEWEREIDNTNLIYTLIESQOQKEK 655  
QY 619 NQELLELDKWSLWVFNFDISKWLWYIK 646  
Db 656 NQELLELDKWSLWVFNFDISKWLWYIK 683  
RESULT 13  
S13288  
env protein - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C;Accession: S13288  
R;O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.  
Nature 348, 69-73, 1990  
A;Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
A;Reference number: S13288; MUID:91043044; PMID:2172833  
A;Accession: S13288  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-854 <OBR>  
C;Superfamily: type B retrovirus env polyprotein  
Query Match 83.9%; Score 2914.5; DB 2; Length 854;  
Best Local Similarity 81.8%; Pred. No. 9.1e-203;  
Matches 561; Conservative 35; Mismatches 45; Indels 45; Gaps 11;



A;Reference number: A42995; MUID:92351552; PMID:1322587

A;Accession: A42995

A;Molecule type: mRNA

A;Residues: 1-861 <SHI>

A;Cross-references: GB:S41266; GB:D01206

C;Genetics:

A;Gene: env

C;Superfamily: type E retrovirus env polyprotein

C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F;1-689/Domain: extracellular #status predicted <EXT>

F;1-33/Domain: signal sequence #status predicted <SIG>

F;17-33/Region: hydrophobic #status predicted

F;34-517/Product: coat protein gp120 #status predicted <CP1>

F;514-517/Region: cleavage processing #status predicted

F;518-861/Product: coat protein gp41 #status predicted <CP2>

F;518-534/Region: hydrophobic #status predicted

F;690-711/Domain: transmembrane #status predicted <TM1>

F;712-861/Domain: intracellular #status predicted <INT>

F;756-772/Region: hydrophobic #status predicted

F;93,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414

Query Match 83.7%; Score 2908.5; DB 1; Length 861;

Best Local Similarity 79.0%; Pred. No. 2.5e-202;

Matches 549; Conservative 42; Mismatches 49; Indels 55; Gaps 8;

QY 1 MRVKGIRKNYQHLW-----RGTTLLGLMLMICSAREKLVWTVYVYGVVWKEATTLFCA 54

Db 1 MRVKEIRKNYQHLWRGIMLRWGTMLGLMLMICSAREQLWTVYVYGVVWKEATTLFCA 60

QY 55 SDAKAYDTEVHNWATHACVPTDPNQEIVLENVTENFNMMKNMVEOMHEDIISLWQS 114

Db 61 SDAKAYDTEAHNWATHACVPTDPNQEIVLVNVTENFNMMKNMVEOMHEDIISLWQS 120

QY 115 LKPCVKLTPLCVTLHCTNLKNAATNKSSNMKMDRGEIKNCSPKVGAG----- 162

Db 121 LKPCVKLTPLCVTLHCTDLRNTNNSSIEEK-KGEIKNCSPNVTNIRDVKQKEYALF 179

QY 163 -----KLINCNTSVITQACPVSFEPPIHYCAPAGFAILKCNCKK 203

Db 180 YKLDVVPIDNDNSTNCTRLISCDTSVITQACPVSFEPPIHYCTPAGFALLKCNCKT 239

QY 204 FNGSGPCTNVSTVOCTHGIRPVVSTQLLNGSLAEQGVWIRSENFDTNKTIIIVQLKESV 263

Db 240 FNGTGPKCNVSTVOCTHGIRPVVSTQLLNGSLAEQGVWIRSENFDTNKTIIIVQLNETV 299

QY 264 EINCTRPNNTKRSITIGPGRAFYATGDIIGDIRQAHNCISGEKNNTLKOIVTKLQAOF 323

Db 300 KINCIRPNNTKRVTPGPRVYTTGDIIGDIRQAHNCISRAEWNTLEQIANKURKQF 359

QY 324 GNKTIIVFKQSGGDPETVMHSFNCGGEFFYCNSTQLFNST-----WNNTIGPNNTNGT 376

Db 360 ENKTIIVFNQSSGGDPETVMHFNFCGGEFFYCDSSQLFNSTHLSNGTWNNGT-GPEN--- 414

QY 377 ITLPCRICKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRDGGKEISNTT----EIF 432

Db 415 ITLPCRICKQIIVNMWQEVGKAMYPPIRGQIRCSNITGLLLTRDGGNTQNNNTSSIEIF 474

QY 433 RPKGDMRDNRSELKYKVKVKEPLGVAPTAKRRVVQREKRAVTIGAMFLGFLGAAGS 492

Db 475 RPKGDMRDNRSELKYKVKVKEPLGVAPTAKRRVVQREKRAVGAVFLGFLGAAGS 534

QY 493 TMGARSITLVQARQLLSGIVQOQNLLRAIEAQHLLQLTVMGIK-LQARVLAVERYLK 551

Db 535 TMGAAVTLTVQARQLLPGIVQOQNLLRAIDAOHLLQLTVMGIKQLQARVLAVERYLK 594

QY 552 DQQLGIWGCCKLICTTAVPMNASNKSOLDQIWNNTWMWEEREIDNTNLIYTLIEE 611

Db 595 DQQLGIWGCCKFICTTAVPMNTSWNSKSNFBINWNTWMWEEREINNTNLIYNLIEE 654

QY 612 SQNQEKNEQELLEDKWSLWNPDISKWLWYIK 646

Db 655 SQNQEKNEQDILLADKWDLSLWNPDISKWLWYIK 689

Search completed: February 25, 2004, 14:19:20

Job time : 20.3842 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:03:35 ; Search time 11.2348 Seconds  
(without alignments)  
2994.034 Million cell updates/sec

Title: US-09-891-609a-2  
Perfect score: 3474  
Sequence: 1 MRVKGIRKNYQLWRGGTLL.....DKWASLWNWFDISKWLWYIK 646

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3422	98.5	847	1 ENV_HV1S1	P19550 human immu
2	3032	87.3	843	1 ENV_HV1Y2	P35961 human immu
3	3002.5	86.4	848	1 ENV_HV1JR	P20871 human immu
4	2997.5	86.3	856	1 ENV_HV1SC	P05878 human immu
5	2984	85.9	855	1 ENV_HV1A2	P03378 human immu
6	2947	84.8	865	1 ENV_HV1RH	P04579 human immu
7	2941.5	84.7	852	1 ENV_HV1BN	P12488 human immu
8	2939.5	84.6	856	1 ENV_HV1B1	P03375 human immu
9	2937.5	84.6	868	1 ENV_HV1C4	P05879 human immu
10	2935	84.5	867	1 ENV_HV1J3	P12489 human immu
11	2933	84.4	861	1 ENV_HV1BR	P03377 human immu
12	2929	84.3	855	1 ENV_HV1OY	P20898 human immu
13	2926.5	84.2	856	1 ENV_HV1PV	P03376 human immu
14	2922.5	84.1	856	1 ENV_HV1H2	P04578 human immu
15	2918	84.0	851	1 ENV_HV1B8	P04582 human immu
16	2913.5	83.9	852	1 ENV_HV1S3	P19549 human immu
17	2908.5	83.7	861	1 ENV_HV1KB	P31819 human immu
18	2904	83.6	856	1 ENV_HV1MN	P05877 human immu
19	2896.5	83.4	856	1 ENV_HV1H3	P04624 human immu
20	2894.5	83.3	856	1 ENV_HV1W1	P31872 human immu
21	2891.5	83.2	856	1 ENV_HV1LW	Q70626 human immu
22	2890	83.2	847	1 ENV_HV1W2	P05880 human immu
23	2885.5	83.1	853	1 ENV_HV1MF	P19551 human immu
24	2685	77.3	863	1 ENV_HV1Z8	P05882 human immu
25	2678.5	77.1	859	1 ENV_HV1M2	P04583 human immu
26	2672	76.9	855	1 ENV_HV1Z6	P04580 human immu
27	2663	76.7	853	1 ENV_HV1Z2	P12487 human immu
28	2632.5	75.8	856	1 ENV_HV1ZH	P05881 human immu
29	2611.5	75.2	846	1 ENV_HV1ND	P18799 human immu
30	2595	74.7	853	1 ENV_HV1EL	P04581 human immu
31	2196	63.2	854	1 ENV_SIVCZ	P17281 Chimpanzee
32	1713.5	49.3	421	1 ENV_HV1N5	P12490 human immu
33	1649.5	47.5	460	1 ENV_HV1Z3	P12491 human immu

RESULT 1				
ENV_HV1S1	ID	ENV_HV1S1	STANDARD;	PRT; 847 AA.
AC	P19550;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
DE	glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).			
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11691;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90347835; PubMed=2384920;			
RA	Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;			
RT	"Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";			
RL	J. Virol. 64:4390-4398 (1990).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M55024; AAA45072.1; --			
DR	PDB; 1OBE; 15-MAY-97.			
DR	HIV; M38428; ENV5SF162.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure.			
KW	SIGNAL	1	29	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	30	502	TRANSMEMBRANE GLYCOPROTEIN.
FT	CHAIN	503	847	BY SIMILARITY.
FT	DISULFID	53	73	BY SIMILARITY.
FT	DISULFID	118	203	BY SIMILARITY.
FT	DISULFID	125	194	BY SIMILARITY.
FT	DISULFID	130	155	BY SIMILARITY.
FT	DISULFID	216	245	BY SIMILARITY.
FT	DISULFID	226	237	BY SIMILARITY.
FT	DISULFID	294	328	BY SIMILARITY.
FT	DISULFID	374	435	BY SIMILARITY.
FT	DISULFID	381	408	BY SIMILARITY.
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	135	135	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	232	232	N-LINKED (GLCNAC. . .) (POTENTIAL).

Q02837 simian immu  
P27977 simian immu  
P18040 human immu  
P27757 simian immu  
P20872 human immu  
P32536 human immu  
P24105 human immu  
P05886 simian immu  
P17755 human immu  
P18094 human immu  
P05883 human immu  
P05884 simian immu

ALIGNMENTS

FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	293	293	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	329	329	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	336	336	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	352	352	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	382	382	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	388	388	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	392	392	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	398	398	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	401	401	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	438	438	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	454	454	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	602	602	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	607	607	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	628	628	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
SQ	SEQUENCE	847 AA;	96135 MM;	0A901317FD7FF2AB	CRC64;
Query Match 98.5%; Score 3422; DB 1; Length 847;					
Best Local Similarity 95.4%; Pred. No. 1.Se-256;					
Matches 643; Conservative 0; Mismatches 3; Indels 28; Gaps 2;					
Qy	1	MRVKGIRKYNQHLWRGGTLLGLMLTCSAVEKLVVTVYGVVWKEATTLFCASDAKAY	60		
Db	1	MRVKGIRKYNQHLWRGGTLLGLMLTCSAVEKLVVTVYGVVWKEATTLFCASDAKAY	60		
Qy	61	DTEVHNWATHACVPTDPNPOEIVLENTVFNFNNKNNVQEHEDIIISLDQSLKPCVK	120		
Db	61	DTEVHNWATHACVPTDPNPOEIVLENTVFNFNNKNNVQEHEDIIISLDQSLKPCVK	120		
Qy	121	LTPCLVTLHCTNLKNAITNTKSSNWKEMDRGEIKNCSFKVGAG-----	162		
Db	121	LTPCLVTLHCTNLKNAITNTKSSNWKEMDRGEIKNCSFKVTTTSIRNKQKEYALFYKLDVV	180		
Qy	163	-----KLINCNTSVITQACPKVSFEPIPHYCAPAGFALTKNDKKFNGSGPCTNV	213		
Db	181	PIDNDNTSYKLLINCNTSVITQACPKVSFEPIPHYCAPAGFALTKNDKKFNGSGPCTNV	240		
Qy	214	STVQCTHGIRPVVSTOLLNGSLAEGVWIRSENFNTDNAKTIIVQLKESVEINCTRPNN	273		
Db	241	STVQCTHGIRPVVSTOLLNGSLAEGVWIRSENFNTDNAKTIIVQLKESVEINCTRPNN	300		
Qy	274	TRKSIITGPGRAFYATGDIIGDIRQAHCNISGEKWNNTLKQIVTKLQAPGKNKTIIVFKQS	333		
Db	301	TRKSIITGPGRAFYATGDIIGDIRQAHCNISGEKWNNTLKQIVTKLQAPGKNKTIIVFKQS	360		
Qy	334	SGGDPVLVHVSFNCGGFFCYNSTLPNSTWNTTIGPNTTNGTITLPCRKQIINRWQEV	393		
Db	361	SGGDPVLVHVSFNCGGFFCYNSTLPNSTWNTTIGPNTTNGTITLPCRKQIINRWQEV	420		
Qy	394	GKAMYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFPFGGDMRDNRSELYKYKV	453		
Db	421	GKAMYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFPFGGDMRDNRSELYKYKV	480		
Qy	454	KIEPLGVAPTKARRVQREKRAVTLGAMPLGFLGAGSTMGARSULTLVQARQLLSGIV	513		
Db	481	KIEPLGVAPTKARRVQREKRAVTLGAMPLGFLGAGSTMGARSULTLVQARQLLSGIV	540		
Qy	514	QQQNLLRAEAOHLLQLTWNGIK-1QARVLAVERYLKDOQLLGGCSGKLICTTAVP	572		
Db	541	QQQNLLRAEAOHLLQLTWNGIK-1QARVLAVERYLKDOQLLGGCSGKLICTTAVP	600		
Qy	573	WNASWSNKSLLDQIWNNTWMEWEREIDNYTNLTITLIEESQNOQKNEQELLEDKWSAL	632		
Db	601	WNASWSNKSLLDQIWNNTWMEWEREIDNYTNLTITLIEESQNOQKNEQELLEDKWSAL	660		
Qy	633	WNWFDISKWLWYIK 646			
Db	661	WNWFDISKWLWYIK 674			

RESULT 2					
ENV_HVLY2					
ID	ENV_HVLY2	STANDARD;	PRT;	843 AA.	
AC	P35961;				
DT	01-JUN-1994 (Rel. 29, Created)				
DT	01-JUN-1994 (Rel. 29, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).				
GN	ENV.				
OS	Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).				
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxID=36377;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=93021387; PubMed=1404605;				
RA	Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H., Shaw G.M.;				
RT	"Complete nucleotide sequence, genome organization, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation."				
RL	J. Virol. 66:6587-6600(1992).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL; M93258; -; NOT_ANNOTATED_CDS.				
DR	PIR; H44001; H44001.				
DR	PDB; 1G9N; 27-DEC-00.				
DR	InterPro; IPR000328; Env GP41.				
DR	InterPro; IPR000777; GP120.				
DR	Pfam; PF00516; GP120; 1.				
DR	Pfam; PF00517; GP41; 1.				
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;				
KW	3D-structure.				
FT	SIGNAL	1	29		
FT	CHAIN	30	489		EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	490	843		TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	738	755		POTENTIAL.
FT	DISULFID	53	73		BY SIMILARITY.
FT	DISULFID	118	201		BY SIMILARITY.
FT	DISULFID	125	192		BY SIMILARITY.
FT	DISULFID	130	155		BY SIMILARITY.
FT	DISULFID	214	243		BY SIMILARITY.
FT	DISULFID	224	235		BY SIMILARITY.
FT	DISULFID	292	326		BY SIMILARITY.
FT	DISULFID	373	432		BY SIMILARITY.
FT	DISULFID	380	405		BY SIMILARITY.
FT	CARBOHYD	87	87		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	129	129		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	135	135		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	138	138		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	154	154		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	158	158		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	184	184		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	193	193		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	230	230		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	237	237		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	258	258		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	272	272		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	285	285		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	291	291		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	297	297		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	327	327		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	351	351		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	381	381		N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	389	389		N-LINKED (GLCNAC. . .) (POTENTIAL) .







```
QY 121 LTPLCVTLHCTNLK-----NATYKSSNWKEMDRGEIKNCSFKVGAG----- 162
Db 121 LTPLCVTLHCTNLK-----NATYKSSNWKEMDRGEIKNCSFKVGAG----- 180
QY 163 -----KLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKFNKSGGPC 210
Db 181 KLDVVPIDNTSYLINCNTSVITQACPKVSFEPIPIHYCA-RWFAILNCNNKPKFNGTGPC 239
QY 211 TNYSTVQCTHGIRPVVSTQLLNGSLAEBGVIRSENFDTNACTIIVQLKESVEINCTRP 270
Db 240 TNYSTVQCTHGIRPVVSTQLLNGSLAEBGVIRSENFDTNACTIIVQLKEAVEINCTRP 299
QY 271 NNTRKSIITGPRAFYATGDIIGDROAHNCISGKWNNTLKQIVTKLOAQAGNKTIIVF 330
Db 300 NNNTTSHITGPRAFYATGDIIGDROAHNCISRAKWNNTLKQIVTKLRDQFENKTIIF 359
QY 331 KQSSGDDPELVMSFNCGGFFCYNSTQLPNSWNNTIGPNNTG--TITLPCRKIQIIN 388
Db 360 NRSOGDDPELVMSFNCGGFFCYNSTQLPSSWNGTEGNNNTGGNDTITLPCRKEIIN 419
QY 389 RWOEVGKAMYAPPIRGQIRCSNITGLLLTRDGGKEISNT----TEIFRPGGDMRDNR 444
Db 420 MWQEVGKAMYAPPIKGQVQKSSNITGLLLTRDGSNGSKSNKENTEIFRPGGDMRDNR 479
QY 445 SELYKVKVIEPLGVAPTAKRVVQREKRAV-TLGAMPLGFLGAGSTMGARSUTLTV 503
Db 480 SELYKVKVIEPLGVAPTAKRVVQREKRAVGTTCGAMPLGFLGAGSTMGATSMTLTV 539
QY 504 QARQLLSGIVQOQNLLRAEAQHLLQLTWGK-LQARVLAVRYLKDQQLLGGWGS 562
Db 540 QARQLLSGIVQOQNLLRAEAQHLLQLTWGKIQARVLAVRYLRDQQLLGGWGS 599
QY 563 GKLICTTAVPNWASWNSKSLDQIWNNTWMWEEREIDNTNLYTLIERSONQOEKNEQ 622
Db 600 GKLICTTVPWNTSWNSKSLDKTGNWNTWMWEEREIDNTNLYTLIERSONQOEKNEQ 659
QY 623 LLELDKWSLWNNFDISKMLWYIK 646
Db 660 LLELDKWSLWNNFNITNMLWYIK 683

RESULT 5
ENV HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shiner S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
RT (ARV-2).";
RL Science 227:484-492(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02007; AAB59882.1; -.
```

```
DR PIR; A03976; VCLJJA2.
DR HIV; K02007; ENVSSF2.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 29
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 208 BY SIMILARITY.
FT DISULFID 125 199 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 221 250 BY SIMILARITY.
FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 299 333 BY SIMILARITY.
FT DISULFID 380 442 BY SIMILARITY.
FT DISULFID 387 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 85.9%; Score 2984; DB 1; Length 855;
Best Local Similarity 82.4%; Pred. No. 1.1e-222;
Matches 563; Conservative 34; Mismatches 48; Indels 38; Gaps 6;
```

```
QY 1 MBVKGIRKNYQHLWGGTLLGLMLTCSAVEKLVTVTVYGVVWKEATTLFCASDAKAY 60
Db 1 MBVKGIRKNYQHLWGGTLLGLMLTCSATEKLVTVTVYGVVWKEATTLFCASDARAY 60
QY 61 DTEVHNWATHACVPTDNPQBIENVTFENFMNKNMVEQMHEDIISLWDSLKPCVK 120
Db 61 DTEVHNWATHACVPTDNPQBIENVTFENFMNKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPLCVTLHCTNLKATNTKSNWKEMDRGEIKNCSFKVGAG----- 162
Db 121 LTPLCVTLHCTNLKATNTKSNWKEMDRGEIKNCSFKVGAG----- 180
QY 163 -----KLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKFNKSG 208
Db 181 PIDNASTTTNTYNYRLIHCNRSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKFNKSG 240
```



```
QY 319 LQAPGKNTIVFQSSGDDPEIWMHSPNCGEFPYCNSTOLFNSWNTTIGPNNTNG--T 376
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 LREQFONKTVFTSSSGDDPEIVLHSPNCGEFPYCNSTOLFNSWNTTIGPNNTNGNDT 420
QY 377 ITLPCRIKQIILNWEQVGMAYAPPIRGQIRCSNITGLLTTDGGKEISNTTEIRPGG 436
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 ITLPCRIKQIILNWEQVGMAYAPPIISGQIKCSNITGLLTTDGGEDFTNTTEIRFLGG 480
QY 437 GMDRNRSLYKYKVKIPLGVAPTKAKRRVVQREKRAV-TLGAMFLGLGAAGSTWG 495
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 GNMEDNRSLYKYKVKIPLGVAPTKAKRRVVQREKRAVGTIGAMFLGLGAAGSTWG 540
QY 496 ARSLTITVQARQLLSGIVQQNNLLRAIEAQHLLQLTWGIIK-LQARVLAVERYLKDQQ 554
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 AGSITLTVQARHLLSGIVQQNNLLRAIEAQHLLQLTWGIIKQLQARVLAVERYLRDQQ 600
QY 555 LLGIMCSGKLICTTAVPNWASNSKSLDOIWNMTWMEWERIDNYNLIYTLIESQN 614
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 LLGIMCSGKLICTTAVPNWASNSKSLNMIWNMTWMEWERIDNYTGIYLLIESQN 660
QY 615 QQKNEQELLELDKASLWNNFDSIKWLWYIK 646
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 661 QQKNEQELLELDKAWLWNNFDTQWLWYIR 692

RESULT 7
ID ENV HV1BN STANDARD; PRT; 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (IBR isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89085613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA Dandekar S.;
RT "Biological and molecular characterization of human immunodeficiency
RT virus (HIV-1BR) from the brain of a patient with progressive
RT dementia.";
RL Virology 168:79-89(1989).
CC -|- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21098; AAA44221.1; -.
DR PIR; A31667; VCLJBR.
DR PDB; 1IM7; 23-OCT-02.
DR HIV; M21098; ENV$BRVA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 507 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 508 852 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
```

Query Match 84.7%; Score 2941.5; DB 1; Length 852;

Best Local Similarity 81.8%; Pred. No. 2.1e-219;

Matches 557; Conservative 40; Mismatches 47; Indels 37; Gaps 8;

QY 1 MRVKGIRKQYQHLWR-GGTLGLMLMICSAVEKLWTVVYGVVWKEATTTLCASDAKA 59

Db 1 MRVKGIRKQYQHLWRGGMWLLGILMICSATDKLWTVVYGVVWKEATTTLCASDAKA 60

QY 60 YDTEVNVWATHACVPTDPNQEIVLENTENFNMKNMKNMVEQMHEDIISLWDQSLKPCV 119

Db 61 YDTEIHNWATHACVPTDPNQEIVLENTENFNMKNMKNMVEQMHEDIISLWDQSLKPCV 120

QY 120 KLTPLCVTLHCTNLKNATWTKSSNWKEMDRGRTKNCSEFKVGAG----- 162

Db 121 KLTPLCVTLNCHDF-NATNATNSGOMEGEGNCNFTTISIRDMOKEYALFYKLDI 179

QY 163 -----KLINCNTSVITQACPKVSFPPIPIHYCAPAGFAILKCNCKKFNKSGPC 210

Db 180 VPIDNDKTWTRYLRILSCNTSVITQACPKVTFPIPIHYCAPAGFAILKCNCKKFNKSGPC 239

QY 211 TNVSTVQCTHGIRPVVSTOLLNGSLAEBGVVIRSENFTDNATKIIVQLKESVEINCTRP 270

Db 240 TNVSTVQCTHGIRPVVSTOLLNGSLAEBGVVIRSENFTNNVKTIIIVQLNESVEINCTRP 299

QY 271 NNNTKRSITIGPCRAFYATGDIIGDIPROACHNCSGKKNWTLKQIVTKLQAOFGNKTIIV 330

Db 300 NNNTKRIITWGPGRVYITTTGIIIGDIRRAHCLNLSRKNWTLKQIVTKURVQFNKTIIV 359

QY 331 KQSSGGDPEIWMHSPNCGEFPYCNSTOLFNSWNTTIGPNNTNGT--ITLPCRIKQII 387

Db 360 NRSNGGDPPEIWMHSPNCGEFPYCNSTOLFNSWNTTIGPNNTNGT--ITLPCRIKQII 418

QY 388 NRWQEVGMAYAPPIRGQIRCSNITGLLTTDGGKEISNT-TEIFRPGGMDRNRWSE 446

|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Db	419	NM005VGMKAMYPPIRGOIKCSSNITGULLLRDGGNNNETTDTTEIFRPGGGNMNRDNWSE	476
Qy	447	LYKYKVKVIEPLGVAPTAKERVVOREKRAVTLGAMFLGFLGAAGSTWGAASLTLTVQAR	506
Db	479	LYKYKVKVIEPLGVAPTAKERVVOREKRAVTLGAMFLGFLGAAGSTWGAASLTLTVQAR	538
Qy	507	QLLSGIVQQNNLLRAIEAQHLLQLTWGIK-LQARVLAVERYLKDQQLGIWGCSGKL	565
Db	539	LLLSGIVQQNNLLMAIEAQHMLLELTWGIKQLQARVLAVERYLKDQQLGIWGCSGKL	598
Qy	566	ICTTAVPNASWSNKSLSLDQIWNMTWMEWEIDNYTNLIVLIESNQOKEKNEELLE	625
Db	599	ICTTAVPNASWSNKSLSLDIWNMTWMEWEIDNYTNLIVLIESNQOKEKNEELLE	658
Qy	626	LDKWASLWNWFDISKWLWYIK	646
Db	659	LDKWASLWNWFDITNLWLYIK	679
RESULT 8			
ENV_HV1B1	ID	ENV_HV1B1	STANDARD; PRT; 856 AA.
AC	ENV_F03375;		
DT	21-JUL-1986	(Rel. 01, Created)	
DT	21-JUL-1986	(Rel. 01, Last sequence update)	
DT	15-JUL-1999	(Rel. 38, Last annotation update)	
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane		
DE	glycoprotein (GP120); Transmembrane glycoprotein (GP41)].		
GN	ENV.		
OS	Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11678;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85111123; PubMed=2578615;		
RA	Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,		
RA	Joseph S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,		
RA	Baumelster K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,		
RA	Lautenberger J.A., Papas T.S., Graybe J., Chang N.T., Gallo R.C.,		
RA	Wong-Staal F.;		
RT	"Complete nucleotide sequence of the AIDS virus, HTLV-III.";		
RL	Nature 313:277-284(1985).		
RN	[2]		
RP	DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.		
RX	MEDLINE=90285159; PubMed=2355006;		
RA	Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,		
RA	Gregory T.J.;		
RT	"Assignment of intrachain disulfide bonds and characterization of		
RT	potential glycosylation sites of the type 1 recombinant human		
RT	immunodeficiency virus envelope glycoprotein (gp120) expressed in		
RT	Chinese hamster ovary cells.";		
RL	J. Biol. Chem. 265:10373-10382(1990).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement">http://www.isb-sib.ch/announcement</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; M15654; AAA44205.1; -		
DR	PIR; A03973; VCLJH3.		
DR	HIV; M15654; ENVSEH102.		
DR	InterPro; IPR000328; Env GP41.		
DR	InterPro; IPR000777; GP120.		
DR	Pfam; PF00516; GP120; 1.		
DR	Pfam; PF00517; GP41; 1.		
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;		
KW	Signal.		
FT	SIGNAL	1	30
FT	CHAIN	31	511
FT	CHAIN	512	856
FT	FT	EXTERIOR MEMBRANE GLYCOPROTEIN.	
FT	FT	TRANSMEMBRANE GLYCOPROTEIN.	

QY 380 PCRIQIINRWQVGMAYPIRQIRCSSNITGLLLTRDGGKBIANTTEIFRPGGDM 439  
 DB 417 PCRIQIINRWQVGMAYPIRQIRCSSNITGLLLTRDGGNS--NNESEIFRPGGDM 475  
 QY 440 RDNWRSELKYKVKIEPLGVAPTKAKRRVQREKRAVTLGAMFLGFLGAGSTMGARS 499  
 DB 476 RDNWRSELKYKVKIEPLGVAPTKAKRRVQREKRAVTLGAMFLGFLGAGSTMGARS 535  
 QY 500 TLTVOARQLLSGIVQOQNNLLRAIERAQHLLQTLVWGKIK-LQARVLAVERYLKDQQLLGI 558  
 DB 536 TLTVOARQLLSGIVQOQNNLLRAIERAQHLLQTLVWGKIKQARILAVERYLKDQQLLGI 595  
 QY 559 WCGSGKLICTTAVPWNASNSKLDQIWNMTMWEEREIDNTNLIYTLIESQNOQEK 618  
 DB 596 WCGSGKLICTTAVPWNASNSKLEQIWNMTMWEEREIDNTNLIYTLIESQNOQEK 655  
 QY 619 NEQELLELDKWSLWNNWEDISKWLWYIK 646  
 DB 656 NEQELLELDKWSLWNNWEDISKWLWYIK 683

RESULT 9  
 ENV\_HV1C4  
 ID ENV\_HV1C4 STANDARD; PRT; 868 AA.  
 AC F05879;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (CDC-451 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11687;  
 RN [1]\_TaxID=11687;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87041461; PubMed=3490666;  
 RA Desai S.M., Kalyanaram V.S., Casey J.M., Srinivasan A.,  
 RA Andersen P.R., Devare S.G.;  
 RT "Molecular cloning and primary nucleotide sequence analysis of a  
 RT distinct human immunodeficiency virus isolate reveal significant  
 RT divergence in its genomic sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384 (1986).  
 RN [2]  
 RP SEQUENCE OF 34-43.  
 RX MEDLINE=90253924; PubMed=2187500;  
 RA Kalyanaram V.S., Rodriguez V., Veronese F., Rahman R., Lusso P.,  
 RA DeVico A.L., Copeland T., Oroszian S., Gallo R.C., Sarngadharan M.G.;  
 RT "Characterization of the secreted, native gp120 and gp160 of the human  
 RL immunodeficiency virus type 1.";  
 RL AIDS Res. Hum. Retroviruses 6:371-380 (1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M13137; AAA44311.1; -  
 DR PIR; C25523; VCLJH4.  
 DR HIV; M13137; ENV5CDC45.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDs; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL. 1 33  
 FT CHAIN 34 522 EXTERIOR MEMBRANE GLYCOPROTEIN (GP120).  
 FT CHAIN 523 868 TRANSMEMBRANE GLYCOPROTEIN.  
 FT

FT DISULFID 55 75 BY SIMILARITY.  
 FT DISULFID 120 216 BY SIMILARITY.  
 FT DISULFID 127 207 BY SIMILARITY.  
 FT DISULFID 132 163 BY SIMILARITY.  
 FT DISULFID 229 258 BY SIMILARITY.  
 FT DISULFID 239 250 BY SIMILARITY.  
 FT DISULFID 307 341 BY SIMILARITY.  
 FT DISULFID 387 456 BY SIMILARITY.  
 FT DISULFID 394 429 BY SIMILARITY.  
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 245 245 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 342 342 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 349 349 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 473 473 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 649 649 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 828 828 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 868 AA; 98698 MW; A1527FC52A6F0C8 CRC64;

Query Match 84.6%; Score 2937.5; DB 1; Length 868;  
 Best Local Similarity 79.5%; Pred. No. 4.4e-219;  
 Matches 556; Conservative 37; Mismatches 47; Indels 59; Gaps 8;  
 QY 1 MRVKGIRKNYQHLWRGGTLLGLMGLMCSAVEKLVWTVYGVVWKEATTLFCASDAKAY 60  
 DB 3 MEAKGIRKNCOHLWRGWTLLGLMGLMCSAAANLWTVYGVVWKEATTLFCASDAKAY 62.  
 QY 61 DTEVHNWATHACVPTDNPQBIIVLENTENFMNKNVQMHEDIISLWQSLKPCVK 120  
 DB 63 DTEAHNWATHACVPTNPNPQBIIVLENTENFMNKNVQMHEDIISLWQSLKPCVK 122  
 QY 121 LPLCLVTLHCT--NLKNATNTKSN---WKMDRGEIKNCSFKVAG----- 162  
 DB 123 LTPCLVTLCTDLNTNTNTTSLSIIVVWQRGKGMKSNFNTTIRDKVQREYALP 182  
 QY 163 -----KLINCNTSVITQACPKVSEPIPIHYCAPAGAILKCNCK 202  
 DB 183 YKLDVEPIDDKNTNTNTNTKYRLINCNTSVITQACPKVSEPIPIHYCTPTGFALLKCNCK 242  
 QY 203 KFGSGPCTNVSTVQCTHGIRPVSTQLLNGSLABEGVIRSENFDTNAKTIIVOLKES 262  
 DB 243 KFGNGPCTNVSTVQCTHGIRPVSTQLLNGSLABEEVIRSENFNTNAKTIIVOLNVS 302  
 QY 263 VEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAHNCNISGEKWNNTLKOIVTKLQAO 322  
 DB 303 VEINCTRPNNTRKRVTLGPRVWTTTGILNIRQAHNCNISRAQWNNTLQOIATTLREQ 362  
 QY 323 FGKNTIIVKQSSGGDPEIVMHSFNCGBGFYCNSTQLFNNTWNTIGPNTNGT----- 376  
 DB 363 FGKNTIIVKQSSGGDPEIVMHSFNCGBGFYCNSTQLFNNTWNTIGPNTNGT----- 376

QY 377 -----ITLPCRRIKQINRWGVGKAMVAPPIRGQIRCSNITGLLLTRDGGKBSINTT 429  
 Db 418 QKDTGDIITLPCRRIKQINRWGVGKAMVAPPIRGQIRCSNITGLLLTRDGGGE-NQTT 476  
 QY 430 EIFRPGGDMRDNRSSELYKYVVKIEPLGVAPTAKRRVQREKRAV-TLGAMFLGFLG 488  
 Db 477 EIFRPGGDMRDNRSSELYKYVVKIEPLGVAPTAKRRVQREKRAVGLGAMFLGFLG 536  
 QY 489 AGSTWGASRLTITVQARQLLSIVQOQNLLRAIQAQHLQLTWGKIK-IQARVLAVE 547  
 Db 537 AAGSTWGATSMALTQARQLLSIVQOQNLLRAIQAQHLQLTWGKIKQARILAVE 596  
 QY 548 RYKDOQLLGWCSGKLCTTAVPNASWSNKSLDQIWNNTWMEWEIENDYNTLIYT 607  
 Db 597 RYKDOQLLGWCSGKLCTTAVPNASWSNKSLDQIWNNTWMEWEIENDYNTLIYT 656  
 QY 608 LIEESQOQEKQOQLLELDKWSLWNWFDISKWLYIK 646  
 Db 657 LIEESQOQEKQOQLLELDKWSLWNWFDISKWLYIK 695

RESULT 10

ENV\_HV1J3 STANDARD; PRT; 867 AA.  
 AC P12489; Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11694;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89352108; PubMed=2669897;  
 RA Komiya N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;  
 RT "Nucleotide sequences of gag and env genes of a Japanese isolate of  
 RT HIV-1 and their expression in bacteria."  
 RL AIDS Res. Hum. Retroviruses 5:411-419(1989).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M21138; AB03526.1; -.  
 DR HIV; M21138; ENVJH3.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 KW Signal.  
 FT CHAIN 1 30  
 FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 517 867 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 217 BY SIMILARITY.  
 FT DISULFID 125 208 BY SIMILARITY.  
 FT DISULFID 130 160 BY SIMILARITY.  
 FT DISULFID 230 259 BY SIMILARITY.  
 FT DISULFID 240 251 BY SIMILARITY.  
 FT DISULFID 308 342 BY SIMILARITY.  
 FT DISULFID 388 457 BY SIMILARITY.  
 FT DISULFID 395 430 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 867 AA; 98399 MW; 5F2310146B8E8680 CRC64;  
 Query Match 84.5%; Score 2935; DB 1; Length 867;  
 Best Local Similarity 80.2%; Pred. No. 6.9e-219;  
 Matches 558; Conservative 38; Mismatches 48; Indels 52; Gaps 7;  
 QY 1 MRVKGIRKQYQHLWRGCTLLGLMLTCSAVEKLVVTVVYGVVWKEATTLFCASDAKAY 60  
 Db 1 MRVKGIRKQYQHLWRGCTLLGLMLTCSAAEQLVVTVVYGVVWKEAATTLFCASDAKAY 60  
 QY 61 DTEVHNWATHACVPTDNPQEIPLVNTNFNMKNMKNMVEOMHEDIISLDQSLKPCVK 120  
 Db 61 DTEVHNWATHACVPTDNPQEVVLENTFNMKNMKNMVEOMHEDIISLDQSLKPCVK 120  
 QY 121 LTPLCVTLHCTN-----LKNATNTKSNWKMDRGEIKNCSFKVAG----- 162  
 Db 121 LTPLCVTLHCTN-----LKNATNTKSNWKMDRGEIKNCSFKVAG----- 162  
 QY 163 -----KLINCSVTITQACPKVSPPIPIHYCAPAGFALKC 199  
 Db 163 -----KLINCSVTITQACPKVSPPIPIHYCAPAGFALKC 199  
 QY 181 KHDVWPINNSTDKNIKNDNSTRYLLISCNSTSVITQACPKISFPIPIHYCAPAGFALIC 240  
 Db 181 KHDVWPINNSTDKNIKNDNSTRYLLISCNSTSVITQACPKISFPIPIHYCAPAGFALIC 240  
 QY 200 NDKKFNKSGPCTNVSTVQCTHGIRPVVSTQLLINGSLAEEGVVIRSENFDTNKTIVQL 259  
 Db 200 NDKKFNKSGPCTNVSTVQCTHGIRPVVSTQLLINGSLAEEGVVIRSENFDTNKTIVQL 259  
 QY 241 NDKKFNKSGPCTNVSTVQCTHGIRPVVSTQLLINGSLAEEGVVIRSENFDTNKTIVQL 300  
 Db 241 NDKKFNKSGPCTNVSTVQCTHGIRPVVSTQLLINGSLAEEGVVIRSENFDTNKTIVQL 300  
 QY 260 KESVEINCTRPNNNTRKSIITGPGAFYATGDIIGDIRQAHNCNISEKNNNTLKQIVTKL 319  
 Db 260 KESVEINCTRPNNNTRKSIITGPGAFYATGDIIGDIRQAHNCNISEKNNNTLKQIVTKL 319  
 QY 301 KEPVINCITRPSKTRRRRHHIGPGAFYTTKQIAGDLRQAHNCINRARNWATLKQIVGKL 360  
 Db 301 KEPVINCITRPSKTRRRRHHIGPGAFYTTKQIAGDLRQAHNCINRARNWATLKQIVGKL 360  
 QY 320 QAQFGNKTIVFKOSSGDDPEIVMHSFNCGGEPFYCNSTOLF-----NSTWNNTIGPNTT 373  
 Db 320 QAQFGNKTIVFKOSSGDDPEIVMHSFNCGGEPFYCNSTOLF-----NSTWNNTIGPNTT 373  
 QY 361 RKQFVNKTIVFNRRSGDPEIVMHSFNCGGEPFYCNSTOLFNSLTNSNTWNTDEGSNT 420  
 Db 361 RKQFVNKTIVFNRRSGDPEIVMHSFNCGGEPFYCNSTOLFNSLTNSNTWNTDEGSNT 420  
 QY 374 NG--TITLPCRRIKQIINRWGVGKAMVAPPIRGQIRCSNITGLLLTRDGGKBSINTTEI 431  
 Db 374 NG--TITLPCRRIKQIINRWGVGKAMVAPPIRGQIRCSNITGLLLTRDGGKBSINTTEI 431  
 QY 421 GGNDDTITLPCRRIKQIINRWGVGKAMVAPPIRGQIRCSNITGLLLTRDGG--DNQNETET 479  
 Db 421 GGNDDTITLPCRRIKQIINRWGVGKAMVAPPIRGQIRCSNITGLLLTRDGG--DNQNETET 479  
 QY 432 FRPGGDMRDNRSSELYKYVVKIEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAAG 491  
 Db 432 FRPGGDMRDNRSSELYKYVVKIEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAAG 491  
 QY 480 FRPGGDMRDNRSSELYKYVVKIEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAAG 539  
 Db 480 FRPGGDMRDNRSSELYKYVVKIEPLGVAPTAKRRVQREKRAVTLGAMFLGFLGAAG 539  
 QY 492 STMGASRLTITVQARQLLSIVQOQNLLRAIQAQHLQLTWGKIK-LOARVLAVERYL 550  
 Db 492 STMGASRLTITVQARQLLSIVQOQNLLRAIQAQHLQLTWGKIK-LOARVLAVERYL 550  
 QY 540 STMGASRLTITVQARQLLSIVQOQNLLRAIQAQHLQLTWGKIK-LOARVLAVERYL 598  
 Db 540 STMGASRLTITVQARQLLSIVQOQNLLRAIQAQHLQLTWGKIK-LOARVLAVERYL 598















**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:12:26 ; Search time 40.8538 Seconds  
(without alignments)  
4989.132 Million cell updates/sec

Title: US-09-891-609a-2  
Perfect score: 3474  
Sequence: 1 MRVKGIRKRYQHLLWRGGTLL.....DKWASLWNWFDISKWLYIK 646

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mmc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_rvirus:  
16: sp\_bacteriap:  
17: sp\_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	3323	95.7	Q8ar21	Q8ar21 human immun
2	3096.5	89.1	Q7SVL4	Q7SVL4 human immun
3	3093.5	89.0	Q7SVL3	Q7SVL3 human immun
4	3092.5	89.0	Q7SVL7	Q7SVL7 human immun
5	3079.5	88.6	Q7SVL5	Q7SVL5 human immun
6	3074	88.5	Q7SVL6	Q7SVL6 human immun
7	3049.5	87.8	Q40222	Q40222 human immun
8	3040.5	87.5	Q9PXW7	Q9PXW7 human immun
9	3035	87.4	Q75760	Q75760 human immun
10	3030	87.2	Q41883	Q41883 human immun
11	3025.5	87.1	Q70150	Q70150 human immun
12	3022.5	87.0	Q92761	Q92761 human immun
13	3022	87.0	O56108	O56108 human immun
14	3020	86.9	O56110	O56110 human immun
15	3017	86.8	Q70679	Q70679 human immun
16	3017	86.8	Q03811	Q03811 human immun

ALIGNMENTS

RESULT 1

Q8AR21	PRELIMINARY;	PRT;	847 AA.
ID	Q8AR21		
AC	Q8AR21;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Envelope glycoprotein.		
GN	ENV.		
OS	Human immunodeficiency virus 1.		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11676;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SHIVSF162P3.2;		
RA	Gao F.;		
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF536757; AAN05642.1; -		
DR	GO; GO:0019031; C: viral envelope; IEA.		
DR	GO; GO:0005198; F: structural molecule activity; IEA.		
DR	InterPro; IPR000328; Env GP41.		
DR	InterPro; IPR000777; GP120.		
DR	Pfam; PF00516; GP120; 1.		
DR	Pfam; PF00517; GP41; 1.		
SQ	SEQUENCE 847 AA; 95879 MW; 9CF863B97B3BB54C CRC64;		
Query Match 95.7%; Score 3323; DB 15; Length 847;			
Best Local Similarity 92.4%; Pred. No. 8.3e-264;			
Matches 623; Conservative 7; Mismatches 16; Indels 28; Gaps 2;			
QY	1	MRVKGIRKRYQHLLWRGGTLLGLMLICSAVEKLVWTVYGVFWKKEATTLFCASDAKAY 60	Q7SVL1 human immun
DB	1	MRVKGIRKRYQHLLWRGGTLLGLMLICSAVEKLVWTVYGVFWKKEATTLFCASDAKAY 60	Q74812 human immun
QY	61	DTEVHNWATHACVPTDPNPQELVLENTENFMNKNVQMHEDIISLDOSLPCVK 120	Q77694 human immun
DB	61	DTEVHNWATHACVPTDPNPQELVLENTENFMNKNVQMHEDIISLDOSLPCVK 120	Q72993 human immun
QY	121	LTPLCVTLHCTNLKNTATNTKSSNWKEMDRGEIKNCSEFKVGAG----- 162	Q78225 human immun





RT "U.S. Human Immunodeficiency Virus Type 1 Epidemic: Date of Origin,  
RT Population History, and Characterization of Early Strains.";  
RL J. Virol. 77:6359-6366(2003).  
DR EMBL: AY247220; AAP37148.1; -.  
SQ SEQUENCE 852 AA; 96680 MW; F8D81102805D998 CRC64;

Query Match 88.6%; Score 3079.5; DB 15; Length 852;  
Best Local Similarity 85.3%; Pred. No. 8.3e-244;  
Matches 580; Conservative 27; Mismatches 38; Indels 35; Gaps 4;

QY 1 MRVKGIRKQYQHLWRGGTLLGLMLICSAVEKLVTVVYGVVPWKAEATTLFCASDAKAY 60  
DB 1 MRVKGIRKQYQHLWRGGTLLGLMLICSAVEKLVTVVYGVVPWKAEATTLFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNWKNMVEQMHEDIISLWQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNWKNMVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPLCVTLHCTNLKNATNTKSSNWKEMDRGEIKNCSPKVGAG----- 162  
DB 121 LTPLCVTLHCTNLKNATNTKSSNWKEMDRGEIKNCSPKVGAG----- 162  
QY 163 -----KLINCNTSVITQACPKVSFEPPIHYCAPAGFAILKCNKKFNGSGPCTNV 213  
DB 163 -----KLINCNTSVITQACPKVSFEPPIHYCAPAGFAILKCNKKFNGSGPCTNV 213  
QY 214 STVQCTHGIRPVVSTOLLNGSLAEGVIRSNFTDNAKTIIVQLKESVEINCTRPNN 273  
DB 241 STVQCTHGIRPVVSTOLLNGSLAEGVIRSNFTDNAKTIIVQLKESVEINCTRPNN 300  
QY 274 TRKSIITGPRAFYATGDIIGDIRQAHNCISGKWNNTLKOIVTKLQAFGNKTIIVFKOS 333  
DB 301 TRKSIITGPRAFYATGDIIGDIRQAHNCISGKWNNTLKOIVTKLQAFGNKTIIVFKOS 360  
QY 334 SGSDPEIVMHSFNCGGEFFYCNSTQFN-----STWNNTIGPNNTNGTITLPCRICKI 387  
DB 361 SGSDPEIVMHSFNCGGEFFYCNSTQFNSTWNASTWNDETSNNTGTTILPCRICKI 420  
QY 388 NRQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSEL 447  
DB 421 NRQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGSN- ENDTEIFRPGGDMRDNRSEL 479  
QY 448 YKVKVVKIEPLGVAPTKAKRRVVQREKRAVTLGAMFLGLGAAGTGAASMTLTQARQ 507  
DB 480 YKVKVVKIEPLGVAPTKAKRRVVQREKRAVTLGAMFLGLGAAGTGAASMTLTQARQ 539  
QY 508 LLSGIVQOQNLLRAEAQOHLQLTVWGIK-LQARVLAVERYLKQOQLLGIWCSGKLI 566  
DB 540 LLSGIVQOQNLLRAEAQOHLQLTVWGIK-LQARVLAVERYLKQOQLLGIWCSGKLI 599  
QY 567 CTTAVPWNASWSNKSLDQIWNNTWMEWEIEDNYNLTLYTLIESQNOQEKNEQEL 626  
DB 600 CTTAVPWNASWSNKSLDQIWNNTWMEWEIEDNYNLTLYTLIESQNOQEKNEQEL 659  
QY 627 DKWASLWNPFDISKWLWYIK 646  
DB 660 DKWASLWNPFDISKWLWYIK 679

RESULT 6  
Q7SVL6 PRELIMINARY; PRT; 853 AA.  
AC Q7SVL6;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=81CA2;  
RX MEDLINE=2628496; PubMed=12743293;  
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairaj A.S.,  
RA Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;  
RT "U.S. Human Immunodeficiency Virus Type 1 Epidemic: Date of Origin,  
RT Population History, and Characterization of Early Strains.";  
RL J. Virol. 77:6359-6366(2003).  
DR EMBL: AY247219; AAP37147.1; -.  
KW Envelope protein.  
SQ SEQUENCE 853 AA; 96545 MW; E3B2830A1261E237 CRC64;

Query Match 88.5%; Score 3074; DB 15; Length 853;  
Best Local Similarity 84.9%; Pred. No. 2.3e-243;  
Matches 590; Conservative 32; Mismatches 31; Indels 40; Gaps 5;

QY 1 MRVKGIRKQYQHLWRGGTLLGLMLICSAVEKLVTVVYGVVPWKAEATTLFCASDAKAY 60  
DB 1 MRVKGIRKQYQHLWRGGTLLGLMLICSAVEKLVTVVYGVVPWKAEATTLFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNWKNMVEQMHEDIISLWQSLKPCVK 120  
DB 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNWKNMVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPLCVTLHCTNLKNATNTKSSNWKEMDRGEIKNCSPKVGAG----- 162  
DB 121 LTPLCVTLHCTNLKNATNTKSSNWKEMDRGEIKNCSPKVGAG----- 162  
QY 163 -----KLINCNTSVITQACPKVSFEPPIHYCAPAGFAILKCNKKFNGSGPCT 211  
DB 181 PIDDNNTSVRLISNTSVITQACPKVSFEPPIHYCAPAGFAILKCNKKFNGSGPCT 240  
QY 212 NVSTVQCTHGIRPVVSTOLLNGSLAEGVIRSNFTDNAKTIIVOLKESVEINCTRPN 271  
DB 241 NVSTVQCTHGIRPVVSTOLLNGSLAEGVIRSNFTDNAKTIIVOLKESVEINCTRPN 300  
QY 272 NNRKSIITGPRAFYATGDIIGDIRQAHNCISGKWNNTLKOIVTKLQAFGNKTIIVFK 331  
DB 301 NNRKSIITGPRAFYATGDIIGDIRQAHNCISGKWNNTLKOIVTKLQAFGNKTIIVFK 360  
QY 332 QSSGDPPEIVMHSFNCGGEFFYCNSTQFNSTWN-----TIGPNNTNGTITLPCRICKI 386  
DB 361 QSSGDPPEIVMHSFNCGGEFFYCNSTQFNSTWNSTWNTEGSGNNTGTTILPCRICKI 420  
QY 387 INRWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGKEISNT--TEIFRPGGDMRDNR 444  
DB 421 VNMWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGG--SNTSGTEIFRPGGDMRDNR 477  
QY 445 SELYKVKVVKIEPLGVAPTKAKRRVVQREKRAVTLGAMFLGLGAAGTGAASMTLTQ 504  
DB 478 SELYKVKVVKIEPLGVAPTKAKRRVVQREKRAVTLGAMFLGLGAAGTGAASMTLTQ 537  
QY 505 ARQLLSGIVQOQNLLRAEAQOHLQLTVWGIK-LQARVLAVERYLKQOQLLGIWCSG 563  
DB 538 ARQLLSGIVQOQNLLRAEAQOHLQLTVWGIK-LQARVLAVERYLKQOQLLGIWCSG 597  
QY 564 KLICTTAVPWNASWSNKSLDQIWNNTWMEWEIEDNYNLTLYTLIESQNOQEKNEQEL 623  
DB 598 KLICTTAVPWNASWSNKSLDQIWNNTWMEWEIEDNYNLTLYTLIESQNOQEKNEQEL 657  
QY 624 LEIDKWSLWNPFDISKWLWYIK 646  
DB 658 LEIDKWSLWNPFDISKWLWYIK 680

RESULT 7  
O40222 PRELIMINARY; PRT; 854 AA.  
AC O40222;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Env polyprotein.



[illegible]

Db 361 SGSDPEIVTHSFNCGGCEFFYCNSTQFNSTWDTGSDNTERITLPCRIKQIINNMQV 420  
Qy 394 GKAMYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNWRSELYKYKV 453  
Db 421 GKAMYAPPPIRGQIRCSNITGLLLTRDGGNNNGTEIFRPVGGDMRDNWRSELYKYKV 480  
Qy 454 KTEPLGVAPTKAKRRVVQREKRAVTLGAMFLGFLGAAGSTMGARSLTLTVQARQLLSGIV 513  
Db 481 KTEPLGVAPTKAKRRVVQREKRAVTLGAMFLGFLGAAGSTMGARSLTLTVQARQLLSGIV 540  
Qy 514 QQQNNLLRAIEAQHLLQLTVWGKIK-LQARVLAVERYLKDQQLLGTWGCCKLICTTAVP 572  
Db 541 QQQNNLLRAIEAQHLLQLTVWGKIKQARVLAVERYLKDQQLLGTWGCCKLICTTAVP 600  
Qy 573 WNASWNSKSLDQIWNNTWMEWEREDNVTNLTLYLIESONQOEKQEBELLELDKWSL 632  
Db 601 WNASWNSKSLD-IWNNTWMEWEREDNVTNLTLYLIESONQOEKQEBELLELDKWSL 659  
Qy 633 WNWFDISKWLWYIK 646  
Db 660 WNWFDITRWLWYIK 673

RESULT 9  
Q75760  
ID Q75760 PRELIMINARY; PRT; 847 AA.  
AC Q75760;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
ON NCBI\_TaxID=11676;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JRFL;  
RX MEDLINE=87206194; PubMed=3646751;  
RA Koyanagi Y., Miles S., Mitsuyasu R.T., Merrill J.E., Vinters H.V.,  
RA Chen I.S.;  
RT "Dual infection of the central nervous system by AIDS viruses with  
RT distinct cellular tropisms.";  
RL Science 236:819-822(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JRFL;  
RX MEDLINE=91043044; PubMed=2172833;  
RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K.,  
RA Zack J.A., Chen I.S.;  
RT "HIV-1 tropism for mononuclear phagocytes can be determined by regions  
RT of gp120 outside the CD4-binding domain.";  
RL Nature 348:69-73(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JRFL;  
RX MEDLINE=92092169; PubMed=1684385;  
RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.;  
RA "HIV-1 env sequence variation in brain tissue of patients with AIDS-  
RT related neurologic disease.";  
RL J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JRFL;  
RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.,  
RA Koyanagi Y., Namazie A., Zhao J., Diagne A., Idler K.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63632; AAB05604.1; -;  
DR PIR; S13289; S13289.  
DR PIR; T09448; T09448.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.

DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005199; F:structural molecule activity; IEA.  
DR InterPro; IPR00328; Env GP41.  
DR InterPro; IPR00777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
SQ AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
KW SEQUENCE 847 AA; 96160 MW; 022D5F24E04FB29F CRC64;  
Query Match 87.4%; Score 3035; DB 15; Length 847;  
Best Local Similarity 85.4%; Pred. No. 3.7e-240;  
Matches 577; Conservative 29; Mismatches 38; Indels 32; Gaps 6;  
Qy 1 MRVKGIRKNYQHLRGGTLGLGLMIMCSAVEKLVWTVVYGVVPWKEATTLTFCASDAKAY 60  
Db 1 MRVKGIRKSQYQLWKGTLGLGLMIMCSAVEKLVWTVVYGVVPWKEATTLTFCASDAKAY 60  
Qy 61 DTEVHNWATHACVPTDNPQBIIVLENTENFMKNNMVEOMHEDIISLWQSLKPCVK 120  
Db 61 DTEVHNWATHACVPTDNPQBIIVLENTENFMKNNMVEOMHEDIISLWQSLKPCVK 120  
Qy 121 LTPLCVTLHCTNLKNAATNTKSSNWKEMDRGEIKNCSFKVGAG----- 162  
Db 121 LTPLCVTLNCKDV-NATNTTNDSEGTMERGEIKNCSFNITTSIRDEVQKEYALFYKLDVV 179  
Qy 163 -----KLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKFNKSGPCTNV 213  
Db 180 PIDNNNTSYRLISCDTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKTENGKGFCKNV 239  
Qy 214 STVQCTHGIRPVVSTQLLNGSLABEGVIRSENFTDNAKTIIVQLKESVEINCTRPNN 273  
Db 240 STVQCTHGIRPVVSTQLLNGSLABEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNN 299  
Qy 274 TRKSIITIGPRAFYATGDIIGDIRQAHCNISGEKNNTLKQIVTKLQADFGNKTIVFKOS 333  
Db 300 TRKSIHIGPRAFYTTGEIIGDIRQAHCNISRAKNWNTLUKQIVKUREFENKTIIVFNHS 359  
Qy 334 SGSDPEIVMHSFNCGGEFFYCNSTQFNSTW-NNTIGPNNNG-TITLPCRIKQIINRWQ 391  
Db 360 SGSDPEIVMHSFNCGGEFFYCNSTQFNSTWNNTEGSNNTGNTITLPCRIKQIINMWQ 419  
Qy 392 EVGKAMYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNWRSELYKYK 451  
Db 420 EVGKAMYAPPPIRGQIRCSNITGLLLTRDGGIN-ENGTEIFRPGGDMRDNWRSELYKYK 478  
Qy 452 VVKIEPLGVAPTKAKRRVVQREKRAVTLGAMFLGFLGAAGSTMGARSLTLTVQARQLLSG 511  
Db 479 VVKIEPLGVAPTKAKRRVVQREKRAVTLGAMFLGFLGAAGSTMGAASMTLTVQARQLLSG 538  
Qy 512 IVQQNNLLRAIEAQHLLQLTVWGKIK-LQARVLAVERYLKDQQLLGTWGCCKLICTTA 570  
Db 539 IVQQNNLLRAIEAQHLLQLTVWGKIKQARVLAVERYLGDQQLLGTWGCCKLICTTA 598  
Qy 571 VPMNASWSNKSLLDQIWNNTWMEWEREDNVTNLTLYLIESONQOEKQEBELLELDKWA 630  
Db 599 VPMNASWSNKSLLDQIWNNTWMEWEREDNVTNLTLYLIESONQOEKQEBELLELDKWA 658  
Qy 631 SLWNWFDISKWLWYIK 646  
Db 659 SLWNWFDITRWLWYIK 674  
RESULT 10  
Q70150  
ID Q70150 PRELIMINARY; PRT; 843 AA.  
AC Q70150;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.



Db 1 MRVKGIRKNCQHLWKWGTMLLGLMICSAAEQLRVTVVYGVVWKEATTLTFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNMKNWVEQMHEDIISLWQSLKPCVK 120  
Db 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNMKNWVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPLCVTLHCTN-LKNATNTKSNWKEMDRGEIKNCSFKY----- 159  
Db 121 LTPLCVTLHCTN-LKNATNTKSNWKEMDRGEIKNCSFKY----- 159  
QY 160 -----GAGKLCNNTSVITACPKVSEPIPIHYCAPAGFAILKCNCKKFNKSGPCTN 212  
Db 181 VPIDNTTTSYRLNLCNNTSVITACPKVSEPIPIHYCAPAGFAILKCNCKKFNKSGPCTN 240  
QY 213 VSTVQCHTGRIPVYSTOLLNGSLAEVGVVIRENFNDNAKTIIVQLKESVEINCTRPNN 272  
Db 241 VSTVQCHTGRIPVYSTOLLNGSLAEVGVVIRENFNDNAKTIIVQLKESVEINCTRPNN 300  
QY 273 NTRKSIITIGPGRFYATGDIIGIRQAHNCNISGEKWNNTLKQIVTKLQAFGNGKTIIVFKQ 332  
Db 301 NTRKSIHIGPGRFYATGDIIGIRQAHNCNISGEKWNNTLKQIVTKLQAFGNGKTIIVFKQ 360  
QY 333 SSGDPEIVMHSFNCGGEFFYCNSTOLFNSWTW--NNTIGPNT--NGTITLPCRIKQ 385  
Db 361 SSGDPEIVMHSFNCGGEFFYCNSTOLFNSWTW--NNTIGPNT--NGTITLPCRIKQ 418  
QY 386 IINRWQEVGKAMYAPPIRGQIRCSNITGLLTRDGGKEISNTTEIFRPGGDMRDNRWS 445  
Db 419 IINRWQEVGKAMYAPPIRGQIRCSNITGLLTRDGGKEISNTTEIFRPGGDMRDNRWS 477  
QY 446 ELYKVKVIEPLGAVPTAKRRVVQREKRAVTLGAMFLGFLGAAGSTWGARSILTVOA 505  
Db 478 ELYKVKVIEPLGAVPTAKRRVVQREKRAVTLGAMFLGFLGAAGSTWGARSILTVOA 537  
QY 506 RQLLSGIVQOQNLLRAIEAQHLLQLTVMGK-LQARVLAVERYLKQDQLLGIWCSGK 564  
Db 538 RQLLSGIVQOQNLLRAIEAQHLLQLTVMGK-LQARVLAVERYLKQDQLLGIWCSGK 597  
QY 565 LICTTAVPNWASNSKSLDOIWNMTWMEWERIDNYTLIYTLIESQOQKNEQLELL 624  
Db 598 LICTTAVPNWASNSKSLDOIWNMTWMEWERIDNYTLIYTLIESQOQKNEQLELL 657  
QY 625 ELDKWASLWNFWDISKWLWYIK 646  
Db 658 ELDKWASLWNFWDISKWLWYIK 679

RESULT 12  
O92761 ID O92761 PRELIMINARY; PRT; 852 AA.  
AC O92761; 01-NOV-1998 (T-EMBLrel. 08, Created)  
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP STRAIN=SPMHS2.9;  
RX MEDLINE=98178716; PubMed=9519894;  
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K., Sheppard W.H.;  
RT "Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in San Francisco Men's Health Study participants.";  
RL AIDS Res. Hum. Retroviruses 14:329-337(1998).  
DR ENBL; AF025750; AAC40588.1; --  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0006096; P:glycolysis; IEA.  
DR InterPro; IPR00328; Env GP41.  
DR InterPro; IPR000173; GAP\_dhdrogenase.  
DR InterPro; IPR000777; GPI20.  
DR Pfam; PF00516; GPI20; 1.  
DR Pfam; PF00517; GP41; 1.  
DR PROSITE; PS00071; GAPDH; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 852 AA; 96545 MW; 3C7780DB06115617 CRC64;

Query Match 87.0%; Score 3022.5; DB 15; Length 852;  
Best Local Similarity 83.1%; Pred. No. 46-239;  
Matches 564; Conservative 37; Mismatches 45; Indels 33; Gaps 3;

QY 1 MRVKGIRKNCQHLWKWGTMLLGLMICSAVEKLVTVVYGVVWKEATTLTFCASDAKAY 60  
Db 1 MRVKGIRKNCQHLWKWGTMLLGLMICSAAGQLVTVVYGVVWKEATTLTFCASDAKAY 60  
QY 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNMKNWVEQMHEDIISLWQSLKPCVK 120  
Db 61 DTEVHNWATHACVPTDPNPQEIIVLENVTENFNMKNWVEQMHEDIISLWQSLKPCVK 120  
QY 121 LTPLCVTLHCTN-LKNATNTKSNWKEMDRGEIKNCSFKVAG----- 162  
Db 121 LTPLCVTLHCTN-LKNATNTKSNWKEMDRGEIKNCSFKVAG----- 162  
QY 163 -----KLCNNTSVITACPKVSEPIPIHYCAPAGFAILKCNCKKFNKSGPCT 211  
Db 181 PIDTNTNTSYRLNLCNNTSVITACPKVSEPIPIHYCAPAGFAILKCNCKKFNKSGPCT 240  
QY 212 NVSTVQCHTGRIPVYSTOLLNGSLAEVGVVIRENFNDNAKTIIVQLKESVEINCTRPN 271  
Db 241 NVSTVQCHTGRIPVYSTOLLNGSLAEVGVVIRENFNDNAKTIIVQLKESVEINCTRPN 300  
QY 272 NTRKSIITIGPGRFYATGDIIGIRQAHNCNISGEKWNNTLKQIVTKLQAFGNGKTIIVFK 331  
Db 301 NTRKSIHIGPGRFYATGDIIGIRQAHNCNISGEKWNNTLKQIVTKLQAFGNGKTIIVFK 360  
QY 332 QSSGDPEIVMHSFNCGGEFFYCNSTOLFNSWTWNTIGPNT--NGTITLPCRIKQIN 388  
Db 361 QSSGDPEIVMHSFNCGGEFFYCNSTOLFNSWTWNTIGPNT--NGTITLPCRIKQIN 420  
QY 389 RWQEVGKAMYAPPIRGQIRCSNITGLLTRDGGKEISNTTEIFRPGGDMRDNRWSELY 448  
Db 421 RWQEVGKAMYAPPIRGQIRCSNITGLLTRDGGKEISNTTEIFRPGGDMRDNRWSELY 480  
QY 449 KYKVKVIEPLGAVPTAKRRVVQREKRAVTLGAMFLGFLGAAGSTWGARSILTVOARQL 508  
Db 481 KYKVKVIEPLGAVPTAKRRVVQREKRAVTLGAMFLGFLGAAGSTWGARSILTVOARQL 540  
QY 509 LSGIVQOQNLLRAIEAQHLLQLTVMGK-LQARVLAVERYLKQDQLLGIWCSGKLC 567  
Db 541 LSGIVQOQNLLRAIEAQHLLQLTVMGK-LQARVLAVERYLKQDQLLGIWCSGKLC 600  
QY 568 TTAVPNWASNSKSLDOIWNMTWMEWERIDNYTLIYTLIESQOQKNEQLELL 627  
Db 601 TTAVPNWASNSKSLDOIWNMTWMEWERIDNYTLIYTLIESQOQKNEQLELL 660  
QY 628 KWASLWNFWDISKWLWYIK 646  
Db 661 KWASLWNFWDISKWLWYIK 679

RESULT 13  
O56108 ID O56108 PRELIMINARY; PRT; 853 AA.  
AC O56108; 01-JUN-1998 (T-EMBLrel. 06, Created)  
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.



```

QY 389 RHOEVGKAMYAPPPIRGQIRCSSNITGLLLTRDGGKEISNTTIFRPGGDMRDNRSLEY 448
DB 421 MHOEVGKAMYAPPKQIRCSSNITGLLLTRDGENS-NNETETFRPGGDMRDNRSLEY 479
QY 449 KYKVKIEPLGVAPTAKRRVVQREKRAVTLGAMFLGFLGAAGSTWGARSLLTIVQARQL 508
DB 480 KYKVKIEPLGVAPTAKRRVVQREKRAVIGIGAVFLGFLGAAGSTWGAAASMTLTVQARLL 539
QY 509 LSGIVQOQNNLRATEAQOHLQLTVWGIK-LQARVLAVERYLKQOQLLGIWGCSGKLC 567
DB 540 LSGIVQOQNNLRATEAQOHLQLTVWGIKQOARVLAVERYLKQOQLLGIWGCSGKLC 599
QY 568 TTAVPNWSNKSLSDOIWNNTWMEWEREIDNYTNLIYTLIEESQOQEKNEQELLEID 627
DB 600 TTVPKTSWSNKSLEKINWNTWMEWDREINNYTSLIYTLIEESQOQEKNEQELLEID 659
QY 628 KWASLWNPFDISKWLWYIK 646
DB 660 KWASLWNPFDITNLWYIK 678

RESULT 15
O70679 PRELIMINARY; PRT; 845 AA.
AC O70679;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL42;
RC MEDLINE=98150884; PubMed=9491920;
RA Graf M., Shao Y., Zhao Q., Seidl T., Kostler J., Wolf H., Wagner R.;
RT "Cloning and characterization of a virtually full-length HIV type 1
RT genome from a subtype B'-Thai strain representing the most prevalent
RT B-clade isolate in China.";
RL AIDS Res. Hum. Retroviruses 14:285-288(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RL42;
RA Graf M., Shao Y., Seidl T., Kostler J., Wolf H., Wagner R.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U71182; AAC05236.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env.GP41.
DR DR; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 845 AA; 95963 MW; 9AF138531C5D7D93 CRC64;

Query Match 86.8%; Score 3017; DB 15; Length 845;
Best Local Similarity 84.0%; Pred. No. 1.1e-238;
Matches 568; Conservative 28; Mismatches 46; Indels 34; Gaps 5;

QY 1 MRVGIKKNYQHWRGTTLLGLMLICSAVEKLWTVYGVVWKEATTLFCASDAKAY 60
DB 1 MRVGIKKNYQHWRGTTLLGLMLICNAENLWTVYGVVWKEATTLFCASDAKAY 60
QY 61 DTEVHNWATHACVPTDPNPOEIVLENVTENFNWKNMVQEMHEDIISLWDSQKPCVK 120
DB 61 DTEVHNWATHACVPTDPNPOEIVLENVTENFNWKNMVQEMHEDIISLWDSQKPCVK 120
QY 121 LTPCLVTLNCTLNKATNTSS- ---MBGGEIKNCSEFNITTSIKTKVKDYALFYKVDVVP 176
DB 163 -----KLNCSNTSVITQACPKVSFPPIPHYCAPAGFALLKCNDDKFKNGSGPCTNVS 214
DB 177 IGNDSTSVRLNCSNTSVITQACPKVSFPPIPHYCTPAGFALLKCNKKFNGTGPCTNVS 236
QY 215 TVQCTHGIRPVVSTOLLNGSLAEGVWIRSNFTDNAKTIIVQLKESVEINCTRPNNNT 274
DB 237 TVQCTHGIRPVVSTOLLNGSLAEEVWIKFNSFTDNARVIVQLNESVEIKCIRPNNNT 296
QY 275 RKSITIGRPARYATGDIIGIROAHCHNISGEKNWNTLQIIVTKLQAOFGNKTIVFKQSS 334
DB 297 RKSIIHLGFKAWYTTGQIIGIROAHCHNLSTKWNNTLQIITKCLREOFGNKTIVFNQSS 356
QY 335 GGDPEIVMHSFNCGGEFFYCNSTQLPNSTWNTTIGPNNNT--NGTITLPCRILQIINRWQE 392
DB 357 GGDPEIVMHSFNCGGEFFYCNSTQLPNSTWNTTIGPNNNT--NGTITLPCRILQIINRWQE 416
QY 393 VGKAMYAPPPIRGQIRCSSNITGLLLTRDGGKEISNTTIFRPGGDMRDNRSLEYKYKV 452
DB 417 VGKAMYAPPPIRGQIRCSSNITGLLLTRDGGNNESKPTETFRPGGDMRDNRSLEYKYKV 476
QY 453 VKIEPLGVAPTAKRRVVQREKRAV-TLGAMFLGFLGAAGSTWGARSLLTIVQARQLLGG 511
DB 477 VKIEPLGVAPTAKRRVVQREKRAVGTIGAMFLGFLGAAGSTWGAAASITLTVQARQLLGG 536
QY 512 IVQOQNNLRATEAQOHLQLTVWGIK-LQARVLAVERYLKQOQLLGIWGCSGKLCITTA 570
DB 537 IVQOQNNLRATEAQOHLQLTVWGIKQOARVLAVERYLKQOQLLGIWGCSGKLCITTA 596
QY 571 VPMNASWSNKSLSDOIWNNTWMEWEREIDNYTNLIYTLIEESQOQEKNEQELLEIDKWA 630
DB 597 VPMNASWSNKSLSHEIWNNTWMEWEREIDNYTREIYTLIEESQOQEKNEQELLEIDKWA 656
QY 631 SLWNPFDISKWLWYIK 646
DB 657 SLWNPFDITNLWYIK 672

Search completed: February 25, 2004, 14:18:35
Job time : 45.8538 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:02:44 ; Search time 52.8474 Seconds  
(without alignments)  
3309.468 Million cell updates/sec

Title: US-09-891-609A-4  
Perfect score: 3327  
Sequence: 1 SAVEKLWTVYGVVWKEA.....DKWASLWNFDISKWLWYIK 619

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:.\*  
1: geneseqp1980s:.\*  
2: geneseqp1990s:.\*  
3: geneseqp2000s:.\*  
4: geneseqp2001s:.\*  
5: geneseqp2002s:.\*  
6: geneseqp2003as:.\*  
7: geneseqp2003bs:.\*  
8: geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3327	100.0	619	5	Aau75156 N-termina
2	3327	100.0	646	5	Aau75155 Modified
3	3275	98.4	842	6	Abr55684 HIV isola
4	3275	98.4	842	7	Adc13218 Protein o
5	3275	98.4	847	3	Aay97073 Variant H
6	3270	98.3	842	6	Abu66565 Human imm
7	3260	98.0	842	5	Abb06211 HIV Env i
8	2942.5	88.4	850	2	Aar67724 gp120 fro
9	2928	88.0	883	4	Aab82761 Ancestral
10	2919	87.7	856	6	Abr55495 Amino aci
11	2908.5	87.4	851	1	Aap80967 HIV prote
12	2876.5	86.5	855	2	Aaw11581 Human imm
13	2876.5	86.5	855	2	Aaw88113 Env prote
14	2862	86.0	643	4	Aab61505 HIV-1 SOS
15	2862	86.0	643	6	Abr57052 HIV-1 JR-
16	2856	85.8	855	2	Aaw53112 ENV prote
17	2856	85.8	855	3	Aay77238 HIV-1 (AT
18	2856	85.8	855	3	Aay77302 HIV-1 (AT
19	2856	85.8	855	6	Abu57553 AIDS asso
20	2856	85.8	855	6	Abu57550 AIDS asso
21	2856	85.8	855	6	Abu63182 Protein #
22	2856	85.8	855	6	Abu63186 Env prote
23	2856	85.8	863	1	Aap61509 Sequence
24	2856	85.8	863	2	Aar29706 env gene
25	2856	85.8	863	5	Aae35790 ARV-2 (9B

ALIGNMENTS

RESULT 1

AAU75156					
ID	AAU75156	standard; protein; 619 AA.			
XX					
AC	AAU75156;				
XX					
DT	08-MAY-2002	(first entry)			
XX					
DE	N-terminal mutant of modified HIV-1 SF162DV2 gp140 envelope protein.				
XX					
KW	Human immunodeficiency virus type 1; HIV-1; gp140 envelope protein;				
KW	second hypervariable region; V2; SF162deltaV2; SF162DV2; AIDS;				
KW	acquired immunodeficiency syndrome; anti-HIV; mutant; mutein; immunogen.				
XX					
OS	Human immunodeficiency virus 1; strain SF162 (clade B).				
XX					
FH	Key	Location/Qualifiers			
FT	Misc-difference 512	/note= "Encoded by CAGCTC"			
XX					
PN	WO200200250-A2.				
XX					
PD	03-JAN-2002.				
XX					
PF	27-JUN-2001; 2001WO-US020483.				
XX					
PR	27-JUN-2000; 2000US-0214608P.				
PR	26-JUN-2001; 2001US-00891609.				
XX					
PA	(AARO-) AARON DIAMOND AIDS RES CENT.				
XX					
PI	Stanatatos L, Barnett S, Shrivastava I;				
XX					
DR	WPI; 2002-130836/17.				
DR	N-PSDB; AAS15499.				
XX					
PT	Immunizing an animal and eliciting an immune response against				
PT	heterologous HIV-1 in an animal, involves administering an immunogen				
PT	comprising modified HIV-1 envelope protein, or DNA or virus encoding the				
XX					
PS	Claim 5; Fig 17; 62pp; English.				
XX					
CC	The present invention relates to methods for immunising an animal,				
CC	preferably human, against heterologous human immunodeficiency virus type				
CC	1 (HIV-1), and eliciting a heterologous immune response to HIV-1 in that				
CC	animal. The method comprises administering an immunogen having at least				





```
Db 268 TRPNNTRKSIITIGPRAFYATGDIIGDIRQAHNCISGEKWNNTLKQIVTKLQAQFGNKT 327
QY 301 IVFKQSSGGDPEIVMHSFNCGGEFFFCNSQTFNSTWNNNTIGPNTNGTITLPCRILKQII 360
Db 328 IVFKQSSGGDPEIVMHSFNCGGEFFFCNSQTFNSTWNNNTIGPNTNGTITLPCRILKQII 397
QY 361 NRQVEGVKAMYPPIRGQIRCSNITGLLLTRDGGKEISNTTIFRPGGDMRDNRSEL 420
Db 388 NRQVEGVKAMYPPIRGQIRCSNITGLLLTRDGGKEISNTTIFRPGGDMRDNRSEL 447
QY 421 YKVKVKIEPLGVAPTAKRVRVQREKRAVTLGAMFLGELGAGSTWGARSLLTVQARQ 480
Db 448 YKVKVKIEPLGVAPTAKRVRVQREKRAVTLGAMFLGELGAGSTWGARSLLTVQARQ 507
QY 481 LLSGIVQQNNLLRAIEAQHLLQLTWGIGLQARVLAVERYLKDQQLLGICSGKILIC 540
Db 508 LLSGIVQQNNLLRAIEAQHLLQLTWGIGLQARVLAVERYLKDQQLLGICSGKILIC 567
QY 541 TTAVPWNASWSKSLDOIWNNTMTMEWEREIDNTNLIYTLIBESQOQEKNEQELLELD 600
Db 568 TTAVPWNASWSKSLDOIWNNTMTMEWEREIDNTNLIYTLIBESQOQEKNEQELLELD 627
QY 601 KWASLWNNWFDISKWLWYIK 619
Db 628 KWASLWNNWFDISKWLWYIK 646

RESULT 3
ABR55684
ID ABR55684 standard; protein; 842 AA.
XX
AC ABR55684;
XX
DT 23-OCT-2003 (revised)
DT 18-AUG-2003 (first entry)
XX
XX HIV isolate SF162 env polypeptide.
XX HIV; gag; nef; prot; tat; rev; vif; vpr; vpu; env; anti-HIV; vaccine;
KW immune response.
XX
XX Human immunodeficiency virus 1.
XX
XX WO2003020876-A2.
XX
XX 13-MAR-2003.
XX
XX 05-JUL-2002; 2002WO-US021342.
XX
XX 31-AUG-2001; 2001US-0316860P.
XX
XX 16-JAN-2002; 2002US-0349728P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Zur Megede J, Barnett SW, Lian Y;
XX
XX WPI; 2003-278761/27.
XX
XX New expression cassettes and polynucleotides encoding HIV Gag, Nef, Prot,
PT Tat, Rev, Vif, Vpr, Vpu, or Env polypeptides, useful for DNA immunization
PT or generating an immune response against HIV in a subject.
XX
XX Example; Fig 2A-C; 214pp; English.
XX
XX The invention relates to an expression cassette comprising a
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Nef,
CC Prot, Tat, Rev, Vif, Vpr, Vpu, or Env polypeptide. The expression
CC cassettes, HIV polypeptides and polynucleotides encoding the HIV
CC polypeptides are useful for DNA immunization or generating an immune
CC response against HIV in a subject. The polynucleotides are also useful
CC for generating packaging cell lines or producing the HIV polypeptides.
CC Sequences ABR55684-688 represent env polypeptides from various HIV
```

```
CC isolates. (Updated on 23-OCT-2003 to standardise OS field)
XX
SQ Sequence 842 AA;
Query Match 98.4%; Score 3275; DB 6; Length 842;
Best Local Similarity 95.2%; Pred. No. 5e-159; 3; Indels 28; Gaps 2;
Matches 616; Conservative 0; Mismatches 0;
QY 1 SAVEKLLWTVVYVYGVVWKEATTTLCFASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
Db 23 SAVEKLLWTVVYVYGVVWKEATTTLCFASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 82
QY 61 VTENFMWKNMNVQEHEDIISLWDSQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 120
Db 83 VTENFMWKNMNVQEHEDIISLWDSQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 142
QY 121 DRGEIKNCSFKVGAG-----KLNCSNTSVITQACPVS 153
Db 143 DRGEIKNCSFKVTTSTRNKNQKEYALFYKLDVVPIDNDNTSYKLINCSNTSVITQACPVS 202
QY 154 FEPPIPIHYCAPAGFAILKCNDDKKFNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAEEG 213
Db 203 FEPPIPIHYCAPAGFAILKCNDDKKFNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAEEG 262
QY 214 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPRAFYATGDIIGDIRQAH 273
Db 263 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPRAFYATGDIIGDIRQAH 322
QY 274 CNISGEKNNTLKQIVTKLQAQFGNKTIVFKOSSGGDPEIVMHSFNCGGEFFFCNSTOLF 333
Db 323 CNISGEKNNTLKQIVTKLQAQFGNKTIVFKOSSGGDPEIVMHSFNCGGEFFFCNSTOLF 382
QY 334 NSTWNTTIGPNTNGTITLPCRILKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRD 393
Db 383 NSTWNTTIGPNTNGTITLPCRILKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRD 442
QY 394 GQKEISNTTIFRPGGDMRDNRSELYKYKVKIEPLGVAPTAKRVRVQREKRAVTLG 453
Db 443 GQKEISNTTIFRPGGDMRDNRSELYKYKVKIEPLGVAPTAKRVRVQREKRAVTLG 502
QY 454 AMPLGFLGAGSTWGARSLLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTWGIGK-L 512
Db 503 AMPLGFLGAGSTWGARSLLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTWGIGKQL 562
QY 513 QARVLAVERYLKDQQLLGICSGKILICTTAVPWNASWSKSLDOIWNNTMTMEWEREID 572
Db 563 QARVLAVERYLKDQQLLGICSGKILICTTAVPWNASWSKSLDOIWNNTMTMEWEREID 622
QY 573 NYTNLIYTLIBESQOQEKNEQELLELDKWASLWNNWFDISKWLWYIK 619
Db 623 NYTNLIYTLIBESQOQEKNEQELLELDKWASLWNNWFDISKWLWYIK 669

RESULT 4
ADCI3218
ID ADCI3218 standard; protein; 842 AA.
XX
AC ADCI3218;
XX
DT 18-DEC-2003 (first entry)
DE Protein of HIV Type C, SF162 SEQ ID NO 2.
XX
XX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV.
XX
XX Human immunodeficiency virus.
OS
XX WO2003004620-A2.
XX
XX 16-JAN-2003.
PD
XX 05-JUL-2002; 2002WO-US021420.
PF
```

XX 05-JUL-2001; 2001US-0303192P.  
PR 21-AUG-2001; 2001US-0316860P.  
PR 16-JAN-2002; 2002US-0349871P.  
XX  
XX (CHIR ) CHIRON CORP.  
XX (UYST-) UNIV STELLENBOSCH.  
XX  
XX Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Renaburg EU;  
XX  
XX WPI; 2003-221593/21.  
XX  
XX New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.  
XX  
XX Disclosure; Fig 2; 301pp; English.  
XX  
XX The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intramuscularly, intranasally,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This sequence represents an HIV Type C related protein of  
CC the invention.  
XX  
XX Sequence 842 AA;  
SQ  
Query Match 98.4%; Score 3275; DB 7; Length 842;  
Best Local Similarity 95.2%; Pred. No. 5e-159;  
Matches 616; Conservative 0; Mismatches 3; Indels 28; Gaps 2;  
QY 1 SAVEKLWVTYYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNQEIVLEN 60  
DB 23 SAVEKLWVTYYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNQEIVLEN 82  
QY 61 VTENFNWKNMVEQMHEDIISLWDSQKPCVKLTPLCVTLHCTNLKNAITNTKSSNWKEM 120  
DB 83 VTENFNWKNMVEQMHEDIISLWDSQKPCVKLTPLCVTLHCTNLKNAITNTKSSNWKEM 142  
QY 121 DRGEIKNCSPKVGAG-----KLNCTNSVITQACPVS 153  
DB 143 DRGEIKNCSPKVTTSIRNKMKEAYLFYKLDVVPIDNDNTSYKLNCTNSVITQACPVS 202  
QY 154 FEPITHYCAPAGFAILKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAEEG 213  
DB 203 FEPITHYCAPAGFAILKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAEEG 262  
QY 214 VVIRSENFDTNAKTIIVQLKESVEINCTRENNTNRKSIITIGPGRAFYATGDIIGDIRQAH 273  
DB 263 VVIRSENFDTNAKTIIVQLKESVEINCTRENNTNRKSIITIGPGRAFYATGDIIGDIRQAH 322  
QY 274 CNTSGEKNWNTLKQIVTKLAQFGNKTIVPKQSGGDPEIVMHSFNCGGEFFVCNSTQLF 333  
DB 323 CNTSGEKNWNTLKQIVTKLAQFGNKTIVPKQSGGDPEIVMHSFNCGGEFFVCNSTQLF 382  
QY 334 NSTWNNITGPNNTGTITLPCRKIQIINRWQEVGKAWYAPPIRGQIRCSSNITGLLLTRD 393  
DB 383 NSTWNNITGPNNTGTITLPCRKIQIINRWQEVGKAWYAPPIRGQIRCSSNITGLLLTRD 442  
QY 394 GGKEISNTTIFRPGGDMRDNRSSELYKYKVVKIEPLGVAPTAKERVVQREKRAVTLG 453  
DB 443 GGKEISNTTIFRPGGDMRDNRSSELYKYKVVKIEPLGVAPTAKERVVQREKRAVTLG 502  
QY 454 AMFLGFLGAAGSTVGARSLLTIVQARQLLSGIVQQQNNLLURATEAQOHLQLTVWGKIL 512  
DB 503 AMFLGFLGAAGSTVGARSLLTIVQARQLLSGIVQQQNNLLURATEAQOHLQLTVWGKIL 562

QY 513 QARVLAVERYLKDQQLLGWCGSKLICTTAVPWNASWSNKSLLDQIWNNTWMEWEREID 572  
DB 563 QARVLAVERYLKDQQLLGWCGSKLICTTAVPWNASWSNKSLLDQIWNNTWMEWEREID 622  
QY 573 NYTNLIYTLIEESQNEQKNEQELLELDKWASLWVWFDISKWLWYIK 619  
DB 623 NYTNLIYTLIEESQNEQKNEQELLELDKWASLWVWFDISKWLWYIK 669  
RESULT 5  
AA97073  
ID AAY97073 standard; protein; 847 AA.  
XX  
XX AAY97073;  
XX AC  
XX 12-SEP-2003 (revised)  
DT 31-OCT-2000 (first entry)  
XX  
XX Variant HIV-1 SF162 Env gp160.  
XX  
XX HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;  
KW CD4 binding region; V1/V2 loop; bridging sheet.  
XX  
XX Human immunodeficiency virus 1; isolate SF162.  
OS  
PN WO200039303-A2.  
XX  
PD 06-JUL-2000.  
XX  
PF 30-DEC-1999; 99WO-US031272.  
XX  
PR 31-DEC-1998; 98US-0114495P.  
PR 29-SEP-1999; 99US-0156670P.  
XX  
XX (CHIR ) CHIRON CORP.  
PA  
PI Barnett S, Hartog K, Martin E;  
XX  
XX WPI; 2000-465745/40.  
XX  
PT Novel modified HIV Env polypeptides useful as immunizing agents and for  
PT preparing a vaccine to elicit an immune response against a broad range of  
PT HIV subtypes.  
XX  
XX Claim 5; Page 115-117; 139pp; English.  
XX  
XX Novel immunogenic modified human immunodeficiency virus (HIV) envelope  
CC (Env) polypeptides having an amino acid deleted or replaced in the region  
CC corresponding to residues 420-436 or 119-123 and 199-210 relative to  
CC isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV  
CC strain SF162, with numbering relative to isolate HXB-2. The Env  
CC polypeptides are modified so as to expose at least part of the CD4  
CC binding region. The modified HIV Env polypeptides, coding polynucleotides  
CC and constructs, further comprising an adjuvant, are used for inducing an  
CC immune response in an individual. The method involves administering a  
CC first composition comprising a polynucleotide encoding the Env  
CC polypeptide in a priming step and administering a second composition  
CC comprising a modified Env polypeptide as a booster in an amount  
CC sufficient to induce an immune response in the individual. The first  
CC and/or second composition further comprises an adjuvant (claimed). The  
CC intracellularly produced Env polypeptides can be used for a number of  
CC diagnostic and therapeutic purposes to determine the presence of reactive  
CC antibodies and/or Env proteins in a biological sample to aid in the  
CC diagnosis of HIV infection or disease status or as measure of response to  
CC immunization. (Updated on 12-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 847 AA;  
Query Match 98.4%; Score 3275; DB 3; Length 847;  
Best Local Similarity 95.2%; Pred. No. 5e-159;  
Matches 616; Conservative 0; Mismatches 3; Indels 28; Gaps 2;  
QY 1 SAVEKLWVTYYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNQEIVLEN 60





CC of the neutralising epitopes in the V2, V4 and C4 domains are variable,  
 CC the amount of variation is highly constrained. This facilitates the  
 CC design of HIV subunit vaccines that can induce antibodies that neutralise  
 CC the most common HIV strains for a given geographic region. This invention  
 CC provides a multivalent gp120 subunit vaccine where the gp120 present in  
 CC the vaccine is from at least two HIV isolates which have different amino  
 CC acid sequences for a neutralising epitope in these regions. (Updated on  
 CC 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise  
 CC OS field)  
 XX  
 SQ Sequence 850 AA;

Query Match 88.4%; Score 2942.5; DB 2; Length 850;  
 Best Local Similarity 85.3%; Pred. No. 4.2e-142;  
 Matches 557; Conservative 27; Mismatches 32; Indels 37; Gaps 5;

QY 1 SAVEKLVTVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60  
 DB |||||  
 QY 28 SAAEKLWTVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 87  
 DB |||||  
 QY 61 VTENFNNKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 120  
 DB |||||  
 QY 88 VTENFNNKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNAATNTSSNGKM 147  
 DB |||||  
 QY 121 DRGEIKNCSPKVGAG-----KLINCNTSVITQACPKVS 153  
 DB :|||  
 QY 148 ERGEIKNCSPNVTTSIRDMKNEAYLFYKLDVVPIQNDNTSYRLISNTSVITQACPKVS 207  
 DB :|||  
 QY 154 FEPPIHYCAPAGFAILKNDKPNFGPCNTVSTVQCTGIRPVVSTOLLNGSLAEEG 213  
 DB |||||  
 QY 208 FEPPIHYCAPAGFAILKNDKPNFGPCNTVSTVQCTGIRPVVSTOLLNGSLAEE 267  
 DB |||||  
 QY 214 VVIRSENFDTNAKTIIVQLKESVEINCTRENNTKRSITIGPGRAFVATDIIIGDIRQAH 273  
 DB |||||  
 QY 268 VVIRSENFDTNAKTIIVQLKESVEINCTRENNTKRSIHGPGRAFVATDIIIGDIRQAH 327  
 DB |||||  
 QY 274 CNISGEKNNTLKQIVTKLQAGFNTKIVPKQSGGDEIIVMHSFNGGEPFFVCNSTQLF 333  
 DB |||||  
 QY 328 CNLSSTKWNNTLKQIVTKLREHF-NKTIIVFNHSSGGDPEIIVMHSFNGGEPFFVCNTTPLF 386  
 DB |||||  
 QY 334 NS-----TNNTIGPNTNGTITLCRIKQIINRWQEVCKAMVAPPIRQIRCSNITG 387  
 DB |||||  
 QY 387 NSTWNTYTYNTNGSNDTGRNITLQCRIKQIINRWQEVCKAMVAPPIRQIRCSNITG 446  
 DB |||||  
 QY 388 LLITDRGKEISNTTIFRPGGDMRDNWRSELYKYKVKVIEPLGVAPTAKKRVVQREK 447  
 DB |||||  
 QY 447 LLITDRGNN--NSETIFRPGGDMRDNWRSELYKYKVKVIEPLGVAPTAKKRVVQREK 504  
 DB |||||  
 QY 448 RAVTLGAMFLGFLGAAGSTWGARSLLTVQARQLLSGIVQQQNNLLRAIEAQHLLQLTV 507  
 DB |||||  
 QY 505 RAVGIGAVFLGFLGAAGSTWGARSLLTVQARQLLSGIVQQQNNLLRAIEAQHLLQLTV 564  
 DB |||||  
 QY 508 WGIK-LQARVLAVERYLKQDLGLGWCSSKLTCTTAVPNASWSNKSLLDQIWNNTWME 566  
 DB |||||  
 QY 565 WGIKQARVLAVERYLKQDLGLGWCSSKLTCTTAVPNASWSNKSLLDQIWNNTWME 624  
 DB |||||  
 QY 567 WEREIDNYTLNLTLEESQNEQKEQLLEDKWLASLNWFDISKWLWYIK 619  
 DB |||||  
 QY 625 WEREIDNYTLNLTLEESQNEQKEQLLEDKWLASLNWFDITKWLWYIK 677  
 DB |||||

RESULT 9  
 AAB82761  
 ID AAB82761 standard; protein; 883 AA.  
 XX  
 AC AAB82761;  
 XX  
 DT 11-SEP-2003 (revised)  
 DT 29-OCT-2001 (first entry)  
 XX  
 DE Ancestral HIV-1 group M, subtype B gp160 protein.  
 XX  
 KW HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.

XX Human immunodeficiency virus 1.  
 OS WO200160838-A2.  
 XX  
 PN 23-AUG-2001.  
 PD  
 XX 16-FEB-2001; 2001WO-US005288.  
 PF  
 XX 18-FEB-2000; 2000US-0183659P.  
 PR  
 XX (UNIW ) UNIV WASHINGTON.  
 PA  
 XX Mullins JI, Rodrigo AG, Learn GH, Li F;  
 PI  
 XX WPI; 2001-536565/59.  
 DR  
 XX N-PSDB; AAH26468.  
 DR  
 XX  
 PT Preparing an ancestral viral amino acid sequence useful as a vaccine  
 PT comprises determining a recent ancestor of a circulating virus by maximum  
 PT likelihood phylogeny analysis.  
 PT  
 XX  
 PS Claim 8; Page 54; 89pp; English.  
 PS  
 XX  
 CC The present sequence is that of an ancestral HIV-1 group M, subtype B  
 CC gp160 (env gene product) sequence. The invention provides compositions  
 CC and methods for determining ancestral viral gene sequences and ancestral  
 CC viral protein sequences for highly diverse viruses, such as HIV-1. The  
 CC methods use samples of circulating viruses to determine an ancestral  
 CC viral sequence by maximum likelihood phylogeny analysis. In the present  
 CC case, the ancestral HIV-1 subtype B env sequence (see AAH26468) was  
 CC determined using 38 subtype B sequences (obtained from 9 different  
 CC countries) and 3 subtype D (outgroup) sequences. The distances between  
 CC this ancestral viral sequence and circulating strains used to determine  
 CC it were on average 12.3% (range: 8.0-21.0%) while the available specimens  
 CC were 17.3% different from each other (range: 13.3-23.2%). Thus, the  
 CC ancestor sequence was, on average, more closely related to any given  
 CC circulating virus than to any other variant. The ancestral gp160 sequence  
 CC included a wide variety of immunogenically active peptides when processed  
 CC for antigen presentation; nearly all known subtype B CTL epitope  
 CC consensus amino acids were represented. Thus, an immunogenic composition  
 CC to this subtype B ancestor protein will elicit broad neutralising  
 CC antibody against HIV-1 isolates of the same subtype, and will also elicit  
 CC a broad cellular response mediated by antigen-specific T-cells. A claimed  
 CC vaccine composition comprises a viral ancestor protein or its immunogenic  
 CC fragment, especially one derived from the HIV-1 group M subtype B gp160  
 CC ancestral protein. (Updated on 11-SEP-2003 to standardise OS field)  
 CC  
 XX  
 SQ Sequence 883 AA;

Query Match 88.0%; Score 2928; DB 4; Length 883;  
 Best Local Similarity 83.3%; Pred. No. 2.4e-141;  
 Matches 569; Conservative 16; Mismatches 34; Indels 64; Gaps 10;

QY 1 SAVEKLVTVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60  
 DB |||||  
 QY 28 SAAEKLWTVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 87  
 DB |||||  
 QY 61 VTENFNNKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 116  
 DB |||||  
 QY 88 VTENFNNKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNAATNTTSSAT 147  
 DB |||||  
 QY 117 -----WKEMDRGEIKNCSPKVGAG----- 135  
 DB :|||  
 QY 148 TWTTSSGGTMEGKEIKNCSPNVTTSIRDMKNEAYLFYKLDVVPIQNDNTNNTNNTS 207  
 DB :|||  
 QY 136 -KLINCNTSVITQACPKVSFEPPIHYCAPAGFAILKNDKPNFGSPCTNTVSTVQCTHG 194  
 DB |||||  
 QY 208 YRLINCNTSVITQACPKVSFEPPIHYCAPAGFAILKNDKPNFGSPCTNTVSTVQCTHG 267  
 DB |||||  
 QY 195 IRPVVSTOLLNGSLAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRENNTKRSITIG 254  
 DB |||||  
 QY 268 IRPVVSTOLLNGSLAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRENNTKRSITIG 327  
 DB |||||

QY 255 PGARFVATGDIIGDIRAOHCNISGEKWNNTLKOIVTKLQAOFGNK--TIVFKQSSGGDPE 312  
DB 328 PGALYATGKIIGDIRAOHCNISRAKWNNTLKOIVTKLREQFGNKKTTIVFKQSSGGDPE 387  
QY 313 IVVHSPNCGGGEFFYCNSTQFNSW--NNTIGPNNT-----NGTTLPCRKQIIN 361  
DB 388 IVVHSPNCGGGEFFYCNSTQFNSWHPFGTGWGNNTERSNNAAADDNDTTLPCRKQIIN 447  
QY 362 RHOEVGKAMYAPPPIRQIRCSSNITGLLLTRDGG--KEISNT-TEIFRPGGDMRDNRWS 418  
DB 448 MHOEVGKAMYAPPISQIRCSSNITGLLLTRDGGNNENTNTDTEIFRPGGDMRDNRWS 507  
QY 419 ELYKYVVKIEPLGVAPTAKRVRVOREKRAV--TLGAMFLGFLGAAGSTMGARSLTLTVQ 477  
DB 508 ELYKYVVKIEPLGVAPTAKRVRVOREKRAVOMLGAMFLGFLGAAGSTMGARSLTLTVQ 567  
QY 478 ARQLSSGIVQQQNNLLRAIEAQHLLQLTVWGIGK-LQARVLAVERYLKDQQLLGWGCSSG 536  
DB 568 ARQLSSGIVQQQNNLLRAIEAQHLLQLTVWGIGK-LQARVLAVERYLKDQQLLGWGCSSG 627  
QY 537 KLICTTAVPWNASWSNKSDDIWNNTWMEWEREIDNYTNLIYTLIEESQOQKNEQEL 596  
DB 628 KLICTTAVPWNASWSNKSDDIWNNTWMEWEREIDNYTNLIYTLIEESQOQKNEQEL 687  
QY 597 LEIDKWSLWNPFDISKWLWYIK 619  
DB 688 LEIDKWSLWNPFDITNLWLWYIK 710

RESULT 10  
ABR55495  
ID ABR55495 standard; protein; 856 AA.

AC ABR55495;  
XX  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Amino acid sequence of a HIV envelope protein.  
XX  
XX Polymorphic marker; host response; HIV; AIDS related virus; vaccine;  
KW hepatitis related virus; HCV; HBV; drug resistance; envelope.  
XX  
XX Human immunodeficiency virus.  
XX  
XX WO2003035097-A1.  
XX  
XX 01-MAY-2003.  
XX  
XX 23-OCT-2002; 2002WO-AU001450.  
XX  
XX 23-OCT-2001; 2001AU-00008425.  
XX  
XX (EPIP-) EPIPOP PTY LTD.  
XX  
XX Mallal S;  
XX  
XX WPI; 2003-449231/42.  
XX  
XX

CC Determining the influence of variation in host genes on the selection of  
PT microorganisms with protein substitutions, comprises typing individuals  
PT of a cohort infected with a microorganism for an intrinsic polymorphic  
PT marker.

PS Claim 22; Page 90-91; 157pp; English.

XX The specification describes a method of determining the influence of  
XX variation in host genes on selection of microorganisms with protein  
XX substitutions. The method comprises typing all individuals of a  
XX population of patients infected with a microorganism for at least one  
XX selected intrinsic polymorphic marker involved in the host response to  
XX the presence of the microorganism. The method is useful for examining  
XX selective pressures confronting a wide range of organisms that exhibit

CC pathogenic traits in a host, such as bacteria, fungi, mycobacterium,  
CC viruses and virus-like particles; for examining microorganisms that have  
CC adapted to evolve rapidly, including HIV and AIDS related viruses and the  
CC hepatitis related viruses such as HCV and HBV. The method is useful for  
CC designing a vaccine to prevent or delay the emergence of drug resistance  
CC in patients treated with a particular drug specific for a microorganism  
CC where the drug affects the replication of the microorganism at the  
CC nucleotide or amino acid level. The present sequence represents a HIV  
CC envelope protein, which is expected to provide optimal cytotoxic T  
CC lymphocyte (CTL) induced therapeutic protection to the cohort examined in  
CC that study  
XX  
SQ Sequence 856 AA;

Query Match 87.7%; Score 2919; DB 6; Length 856;  
Best Local Similarity 85.1%; Pred. No. 6.7e-141;  
Matches 559; Conservative 25; Mismatches 33; Indels 40; Gaps 9;

QY 1 SAVEKLWVTYYGYVPVWKEATTTLCASDAKAYDTEVHNVWATHACVPTDPNPQEVLEN 60  
DB 29 SATEKLWVTYYGYVPVWKEATTTLCASDAKAYDTEVHNVWATHACVPTDPNPQEVLEN 88  
QY 61 VTENFMWKNVVEQMHEDIISLDQSLKPCVKLPLCVTLHCTNLKNATNT-KSSNWKE 119  
DB 89 VTENFMWKNVVEQMHEDIISLDQSLKPCVKLPLCVTLHCTNLKNATNTNTSGSNN 148  
QY 120 MDRGEIKNCSFKVGAG-----KLINCNTSVITQACPKV 152  
DB 149 MEKGEIKNCSFNITTSIRDQKQKVALFYKLDVVPIDNDNTSVRLISCTSVITQACPKV 208  
QY 153 SPEPIPIHYCAPAGFAILKCNCKKNGSGPCTNVSTVQCTHGRPVVSTQLLNGSLAE 212  
DB 209 SPEPIPIHYCAPAGFAILKCNCKKNGTGTCTNVSTVQCTHGRPVVSTQLLNGSLAE 268  
QY 213 GVIRSENFTDNAKTIIVQLKESVEINCTRPNNTRKSI--IGPRAFYATGDIIGDIR 270  
DB 269 EVVIRSENFTNNAKTIIVQLNESVEINCTRPNNTRKSIHIGPRAFYATGE-IGDIR 327  
QY 271 QAHCNISGEKWNNTLKOIVTKLQAOFG-NKTIIVFKOSSGGDPEIVHSPNCGGGEFFYCN 329  
DB 328 QAHCNISRAEWNNTLKOIVKUREQFGKNTIVFNOSGGDPEIVHSPNCGGGEFFYCN 387  
QY 330 TQLFNSWNTN--NTIGPNNTNG--TITLPCRKQIINRWQEVGKAMYAPPPIRQIRCSS 383  
DB 388 TQLFNSWNTNNTWNTESNTEGNETITLPCRKQIINRWQEVGKAMYAPPPIRQIRCSS 447  
QY 384 NITGLLLTRDGGKBIISNTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKRVRV 443  
DB 448 NITGLLLTRDGGNN-NKTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKRVRV 506  
QY 444 QREKRAVTIGAMFLGFLGAAGSTMGARSLTLTVQARQLSSGIVQQQNNLLRAIEAQHLL 503  
DB 507 QREKRAVGIGAMFLGFLGAAGSTMGAASITLTVQARQLSSGIVQQQNNLLRAIEAQHLL 566  
QY 504 QLTWVGIGK-LQARVLAVERYLKDQQLLGWGCSSGKLICTTAVPWNASWSNKSDDIWN 562  
DB 567 QLTWVGIGK-LQARVLAVERYLKDQQLLGWGCSSGKLICTTAVPWNASWSNKSDDIWN 626  
QY 563 TWMEWEREIDNYTNLIYTLIEESQOQKNEQELLELDKWSLWNPFDISKWLWYIK 619  
DB 627 TWMEWEREIDNYTNLIYTLIEESQOQKNEQELLELDKWSLWNPFDISKWLWYIK 683

RESULT 11  
AAP80967  
ID AAP80967 standard; protein; 851 AA.

XX AAP80967;  
AC AAP80967;  
XX  
DT 24-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 15-NOV-1990 (first entry)  
XX

```
DE HIV protein HT6.
XX HIV; HT6; gp160; envelope protein; RF; AIDS.
XX Human immunodeficiency virus; variant RF.
XX EP272858-A.
XX 29-JUN-1988.
XX 14-DEC-1987; 87EP-00310967.
XX 15-DEC-1986; 86US-00941111.
XX 31-AUG-1987; 87US-00091481.
XX (REPK ) REPLIGEN CORP.
XX Rusche J, Lynn D, Carson H, Putney S, Jellis CL;
XX WPI; 1988-176944/26.
XX N-PSDB; AAN80948.
XX Prodn. of recombinant HIV envelope proteins in insect cells - useful as
XX vaccine against AIDS and for diagnosis and therapy.
XX Disclosure; Page ?; 4pp; English.
XX The sequence is the result of cloning a hybrid envelope gene from HIV
XX variants BH10 and RF. A central portion of the RF gene was used, the rest
XX being from te distantly related variant BH10. The resulting clone, pACHT6
XX produces a hybrid gp 160 envelope protein with novel immunological and
XX antigenic characteristics. It may be used to as a vaccine and for
XX diagnosis and therapy of AIDS. See also AAP80966. (Updated on 25-MAR-2003
XX to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)
XX SQ Sequence 851 AA;

Query Match 87.4%; Score 2908.5; DB 1; Length 851;
Best Local Similarity 83.8%; Pred. No. 2.3e-140;
Matches 545; Conservative 39; Mismatches 36; Indels 31; Gaps 5;

QY 1 SAVEKLVTVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
DB 29 SATEKLVTVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 88
QY 61 VTENFNWKNMVEQMHEDIISLDOSLKPVCVLTPLCVTLHCTNLKNAATKSSNKE- 119
DB 89 VTENFNWKNMVEQMHEDIISLDOSLKPVCVLTPLCVTLHCTNLKNAATKSSNKE- 148
QY 120 MRGEIKNCSFKVGA---GK-----LINCNTSVITQACPKV 152
DB 149 MEKGEIKNCSFNMTSIRGKVQKEYAFFYKLDIIPIDNDTTSVTLTSCNTSVITQACPKV 208
QY 153 SFEPPIHYCAPAGFAILKNDKFKNGSGCTVSTVQCTHGRPVVSTOLLNGSLAE 212
DB 209 SFEPPIHYCAPAGFAILKNNKTFTNGTGTCTVSTVQCTHGRPVVSTOLLNGSLAE 268
QY 213 GVVRSENFTDNAKTIIVQLKESVEINCTPNNTRKSIITIGPGRFYATGDIIGDIRQA 272
DB 269 EVVRSENFTDVKTIIVQLNASVQINCTPNNTRKSIITIGPGRFYATGDIIGDIRKA 328
QY 273 HCNISGEKNNTLKQIVTKLQAFGNKTIIVFKOSSGGDPPIVMHFNCGGEPFFYCNSTQL 332
DB 329 HCNLSRAQWNNTLKQIVTKLREQFDNKTIVFTSSGGDPPIVLHFNCGGEPFFYCNSTQL 388
QY 333 FNSTWNTIGPNNTNG--TITLPCRKIQIINRVOEKGKAMYAPPRIQIRCSNITGLLL 390
DB 389 FNSTWNTSTEGNNTGGNDTITLPCRKIQIINRVOEKGKAMYAPPISQIKICISNITGLLL 448
QY 391 TRGGKEISNTTIFRPGGDMRNRSELYKVKVYKIEPLGYAPTKAKRRVVQREKRAV 450
DB 449 TRDGGEDTNTTIFRPGGDMRNRSELYKVKVYKIEPLGYAPTKAKRRVVQREKRAV 508
```

```
QY 451 TLGAMFLGFLGAGSTMGARSLLTAVQARQLLSGIVQOQNLLRLRAEQHLLQLTVMGI 510
DB 509 GIGALFLGFLGAGSTMGASMTLTAVQARQLLSGIVQOQNLLRLRAEQHLLQLTVMGI 568
QY 511 K-LQARVLAVERYLKDOQLLGWCSGKLCITTAVPNNAWSNKSLSDOINNNNTWMEWER 569
DB 569 KQLQARILAVERYLKDOQLLGWCSGKLCITTAVPNNAWSNKSLSQEIWNNTWMEWDR 628
QY 570 EIDNTNLTLYTIEESQOQKNEQELLELDKWSLWNNWFDISKWLWYIK 619
DB 629 EINNTSLIHSLSIEESQOQKNEQELLELDKWSLWNNWFNITNLWYIK 678

RESULT 12
AAW11581
ID AAW11581 standard; protein; 855 AA.
XX
AC AAW11581;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 25-MAR-1997 (first entry)
XX
DE Human Immunodeficiency Virus-1 strain BA-L envelope protein.
XX
KW Acquired immune deficiency syndrome; AIDS; envelope protein; env gene;
KW vaccine.
XX
OS Human immunodeficiency virus 1; (strain BA-L).
XX
PN US5576000-A.
XX
PD 19-NOV-1996.
XX
PF 15-FEB-1995; 95US-00388809.
XX
PR 17-OCT-1990; 90US-00599491.
PR 25-FEB-1993; 93US-00022835.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Gallo RC, Franchini G, Lori FC, Popovic M, Reitz MS, Gartner S;
PI Markham PD;
XX
DR WPI; 1997-011206/01.
DR N-PSDB; AAT58551.
XX
XX New isolated envelope protein of HIV-1 strain BA-L and recombinant
PT equivalents - useful as immunogens for vaccines and antibody prodn.,
PT typical of US clinical isolates.
XX
PS Claim 1; Fig 9; 86pp; English.
XX
CC A HindIII fragment of unintegrated viral DNA representing the BA-L genome
CC was cloned into lambda phage Charon 28 DNA from total DNA of peripheral
CC blood lymphocytes infected with and producing HIV-1(BA-L). A positive
CC clone was selected by hybridisation using a HIV-1 env probe. This clone,
CC designated BA-L1, was found to contain the entire gene for the envelope
CC protein on a 2.8 kb HindIII-XbaI fragment and a 0.4 kb EcoRI-HindIII
CC fragment. When cloned together these fragments comprise the env gene, as
CC well as the coding regions for rev and the rev-responsive element of env,
CC both necessary for efficient expression in eukaryotic cells. The claimed
CC recombinantly produced envelope protein can be used as an immunogen for
CC raising antibodies against HIV. (Updated on 25-MAR-2003 to correct PF
CC field.) (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 855 AA;

Query Match 86.5%; Score 2876.5; DB 2; Length 855;
Best Local Similarity 82.7%; Pred. No. 9.7e-139;
Matches 542; Conservative 33; Mismatches 43; Indels 37; Gaps 5;

QY 1 SAVEKLVTVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
```

```

Db      28 NAEELKLVTVYGVVPWKEATTLFCASDRKAYDEVHNVWATHACVPTDPNPQVEVLKN 87
Qy      61 VTENFNWKNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLHCTNKNATNTKSSNWKEM 120
Db      88 VTENFNWKNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLHCTNKNATNTKSSNWKEM 147
Qy     121 DR-----GEIKNSFKVGAG-----KLINCNTSVITQ 147
Db     148 SRGMVGGGEMKNSFNITNIRKQVKEVALFYKLDIAPIDNNNNRYRLISNTSVITQ 207
Qy     148 ACPKVSFEPIPIHYCAPAGFAILLKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQLLNG 207
Db     208 ACPKVSFEPIPIHYCAPAGFAILLKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQLLNG 267
Qy     208 SLAEEGVIRSENFTDNTAKTIIVQLKESVEINCTRNPNNTKRSITIGPGRAFATDGIIG 267
Db     268 SLAEEVWIRSANFADNAKVIIVQLNESVEINCTRNPNNTKRSIHIGPGRAFATDGIIG 327
Qy     268 DIRQAHCHNLSRAKNDTLNKRIVIKLREQFGNKTIVFKHSSGGDPEIVTHSFNCGGEFFYC 327
Db     328 DIRQAHCHNLSRAKNDTLNKRIVIKLREQFGNKTIVFKHSSGGDPEIVTHSFNCGGEFFYC 387
Qy     328 NSTQLENSTWNNTIGPNNT--NGTITLPCRKIQIINRWQEVGKAMVAPPPIRGQIRCSSNI 385
Db     388 NSTQLENSTWNNTIGPNNT--NGTITLPCRKIQIINRWQEVGKAMVAPPPIRGQIRCSSNI 447
Qy     386 TGLLLTRDGGKEISNTTEIPRPGGDMRDNWRSELYKYVKVIEPLGVAPTAKRRVQR 445
Db     448 TGLLLTRDGGPE-DNKTEVPRPGGDMRDNWRSELYKYVKVIEPLGVAPTAKRRVQR 506
Qy     446 EKRAVTLGAMFLGFLGAGSTMGARSITLTQARQLLSGIVQOQNLLRAIEAQOHLQL 505
Db     507 EKRAVGIGAVFLGFLGAGSTMGARSITLTQARQLLSGIVQOQNLLRAIEAQOHLQL 566
Qy     506 TVMGIK-LOARVLAVERYLKDQQLLGIWGSCKLICTTAVPMNASWSNKSILDKWASLWFDISKWLWYIK 619
Db     567 TVMGIKLOARVLAVERYLKDQQLLGIWGSCKLICTTAVPMNASWSNKSILDKWASLWFDISKWLWYIK 681
Qy     565 MEWERIDNNTNLIYTLIEESQKQEKQELLELDKWLWASLWFDISKWLWYIK 619
Db     627 IEWDREINNTYTSIYSLIEESQKQEKQELLELDKWLWASLWFDISKWLWYIK 681

RESULT 13
ID      AA#88113
AC      AA#88113 standard; protein; 855 AA.
XX      AA#88113;
XX      17-OCT-2003 (revised)
DT      09-APR-1999 (first entry)
XX      XX
DE      Env protein of the BA-L strain of Human immunodeficiency virus type 1.
XX      XX
KW      HIV-1; HIV-1 strain BA-L; env protein; vaccine; immunotherapy;
KW      HIV infection; immunogen; HIV-1 diagnosis.
XX      XX
OS      Human immunodeficiency virus 1.
XX      XX
PH      Key Location/Qualifiers
FT      Misc-difference 11 /note= "Gln encoded by CG"
XX      XX
XX      US5869313-A.
XX      09-FEB-1999.
XX      14-MAY-1996; 96US-00647714.
XX      17-OCT-1990; 90US-00599491.
XX      25-FEB-1993; 93US-00022835.
XX      15-FEB-1995; 95US-00388809.

```

```

XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      Gallo RC, Franchini G, Reitz MS, Lori FC, Popovic M, Gartner S;
XX      Markham PD;
XX      WPI; 1999-152779/13.
XX      N-PSDB; AAX04767.
XX      DNA encoding env protein of the human immune deficiency virus isolate BA-
XX      L - useful for producing protein for use in vaccines, as assay reagent
XX      and to generate antibodies.
XX      Example 1; Fig 9A-C; 87pp; English.
XX      The present sequence represents the envelope protein of the BA-L (ATCC
XX      40890) strain of Human immunodeficiency virus type 1 (HIV-1) strain MN-
XX      ST1. BA-L is more typical of United States isolates of HIV-1 than
XX      previously known strains. Recombinant, complete env protein of the BA-L
XX      strain is used as a vaccine component and for immunotherapy of existing
XX      HIV infections, to detect HIV-specific antibodies, e.g. in donated blood,
XX      and as an immunogen to raise specific antibodies, for HIV-1 diagnosis.
XX      (Updated on 17-OCT-2003 to standardise OS field)
XX      SQ      Sequence 855 AA;
XX      Query Match      86.5%; Score 2876.5; DB 2; Length 855;
XX      Best Local Similarity 82.7%; Pred. No. 9.7e-139;
XX      Matches 542; Conservative 33; Mismatches 43; Indels 37; Gaps 5;
Qy      1 SAVEKLWTVYGVVPWKEATTLFCASDAKAYDEVHNVWATHACVPTDPNPQVEIEN 60
Db      28 NAEELKLVTVYGVVPWKEATTLFCASDRKAYDEVHNVWATHACVPTDPNPQVEVLKN 87
Qy      61 VTENFNWKNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLHCTNKNATNTKSSNWKEM 120
Db      88 VTENFNWKNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLHCTNKNATNTKSSNWKEM 147
Qy     121 DR-----GEIKNSFKVGAG-----KLINCNTSVITQ 147
Db     148 SRGMVGGGEMKNSFNITNIRKQVKEVALFYKLDIAPIDNNNNRYRLISNTSVITQ 207
Qy     148 ACPKVSFEPIPIHYCAPAGFAILLKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQLLNG 207
Db     208 ACPKVSFEPIPIHYCAPAGFAILLKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQLLNG 267
Qy     208 SLAEEGVIRSENFTDNTAKTIIVQLKESVEINCTRNPNNTKRSITIGPGRAFATDGIIG 267
Db     268 SLAEEVWIRSANFADNAKVIIVQLNESVEINCTRNPNNTKRSIHIGPGRAFATDGIIG 327
Qy     268 DIRQAHCHNLSRAKNDTLNKRIVIKLREQFGNKTIVFKHSSGGDPEIVTHSFNCGGEFFYC 327
Db     328 DIRQAHCHNLSRAKNDTLNKRIVIKLREQFGNKTIVFKHSSGGDPEIVTHSFNCGGEFFYC 387
Qy     328 NSTQLENSTWNNTIGPNNT--NGTITLPCRKIQIINRWQEVGKAMVAPPPIRGQIRCSSNI 385
Db     388 NSTQLENSTWNNTIGPNNT--NGTITLPCRKIQIINRWQEVGKAMVAPPPIRGQIRCSSNI 447
Qy     386 TGLLLTRDGGKEISNTTEIPRPGGDMRDNWRSELYKYVKVIEPLGVAPTAKRRVQR 445
Db     448 TGLLLTRDGGPE-DNKTEVPRPGGDMRDNWRSELYKYVKVIEPLGVAPTAKRRVQR 506
Qy     446 EKRAVTLGAMFLGFLGAGSTMGARSITLTQARQLLSGIVQOQNLLRAIEAQOHLQL 505
Db     507 EKRAVGIGAVFLGFLGAGSTMGARSITLTQARQLLSGIVQOQNLLRAIEAQOHLQL 566
Qy     506 TVMGIK-LOARVLAVERYLKDQQLLGIWGSCKLICTTAVPMNASWSNKSILDKWASLWFDISKWLWYIK 619
Db     567 TVMGIKLOARVLAVERYLKDQQLLGIWGSCKLICTTAVPMNASWSNKSILDKWASLWFDISKWLWYIK 681
Qy     565 MEWERIDNNTNLIYTLIEESQKQEKQELLELDKWLWASLWFDISKWLWYIK 619
Db     627 IEWDREINNTYTSIYSLIEESQKQEKQELLELDKWLWASLWFDISKWLWYIK 681

```



RESULT 14	
ID	AAB61505 standard; protein; 643 AA.
XX	
AC	AAB61505;
XX	
DT	11-SEP-2003 (revised)
DT	05-APR-2001 (first entry)
XX	
DE	HIV-1 SOS gp140 glycoprotein.
XX	
KW	gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp41;
KW	gp140.
XX	
OS	Human immunodeficiency virus 1.
XX	
PN	WO200100648-A1.
PD	04-JAN-2001.
XX	
XX	23-JUN-2000; 2000WO-US017267.
XX	
XX	25-JUN-1999; 99US-00340992.
XX	
PA	(PROG-) PROGENICS PHARM INC.
PA	(AARO-) AARON DIAMOND AIDS RES CENT.
XX	
PI	Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;
XX	
DR	WPI; 2001-122993/13.
DR	N-PSDB; AAF28581.
XX	
PT	New viral envelope proteins, useful for producing vaccines to treat human
PT	immunodeficiency virus-1 infections, comprises amino acid sequence
PT	mutations such that viral transmembrane-surface protein complex is more
PT	stable.
XX	
PS	Disclosure; Fig 13; 109pp; English.
XX	
CC	The present invention relates to a viral envelope protein. The viral
CC	envelope protein comprises a viral surface protein (e.g. glycoprotein
CC	gp120) and a corresponding viral transmembrane protein (e.g. gp41), in
CC	which the viral envelope protein contains one or more amino acid sequence
CC	mutations that enhance the stability of the complex formed between the
CC	viral surface and transmembrane proteins. The viral envelope protein can
CC	be used in the treatment of viral infection e.g. HIV-1 infection. The
CC	present sequence is HIV-1 SOS gp140, which was used in the present
CC	invention. (Updated on 11-SEP-2003 to standardise OS field)
XX	
SQ	Sequence 643 AA;
Query Match 86.0%; Score 2862; DB 4; Length 643;	
Best Local Similarity 84.5%; Pred. No. 4.2e-138;	
Matches 545; Conservative 25; Mismatches 43; Indels 32; Gaps 6;	
QY	3 VEKLWTVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDNPQBIVLNVT 62
DB	1 VEKLWTVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDNPQBIVLNVT 60
QY	63 ENFNWKNVNNVEQWHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEMDR 122
DB	61 EHENWKNVNNVEQWHEDIISLWQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTMER 119
QY	123 GEIKNGSKFKVGAG-----KLNKNTSVITQACPKVSFE 155
DB	120 GEIKNGSKFNITTSIRDEVEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179
QY	156 PIPHYCAPAGFAILKCNCKKNGSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEGVV 215
DB	180 PIPHYCAPAGFAILKCNCKNTKNGKPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVV 239
QY	216 IRSENFDTNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPRAFYATGDIIGDIRQAHCN 275
DB	240 IRSDFNTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTTIGDIRQAHCN 299
QY	276 ISGEKWNNTLKQIVTKLQAFQGNKTIIVFKOSSGGDPEIVHMSFNCGGEFFPCNSTQLFNS 335
DB	300 ISRAKWNNTLKQIVTKLQAFQGNKTIIVFNHSSGGDPEIVHMSFNCGGEFFPCNSTQLFNS 359
QY	336 TW-NNTIGPNTNG-TITLPCRIKQIINRWQEVGKAMYAPPPIRGQIRCSNITGLLITRD 393
DB	360 TWNNTEGSGNTEGNTITLPCRIKQIINRWQEVGKAMYAPPPIRGQIRCSNITGLLITRD 419
QY	394 GGEKEINTTEIFRPGGDMRDNRSELYKYKVVVKIEPLGVAPTKARRVVQREKRAVTLG 453
DB	420 GGIN-ENGTEIFRPGGDMRDNRSELYKYKVVVKIEPLGVAPTKARRVVQREKRAVGIG 478
QY	454 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQQNNLLRAEAQOHLLOLTVWGTG-L 512
DB	479 AVFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQNNLLRAEAQOHLLOLTVWGTG-L 538
QY	513 QARVLAVERYLKDQQLLGWCGSKLICCTTAVPNASWSNKSLLDQIWNNTNMEWEREID 572
DB	539 QARVLAVERYLGDQQLLGWCGSKLICCTTAVPNASWSNKSLLDRIWNNTNMEWEREID 598
QY	573 NYTNLIYTLIESQNOQEKNEQELLELDKWLASLWNNFDISKWLWY 617
DB	599 NYTSEIYTLIESQNOQEKNEQELLELDKWLASLWNNFDTNWLWY 643
RESULT 15	
ABR57052	
ID	ABR57052 standard; protein; 643 AA.
XX	
AC	ABR57052;
XX	
DT	23-OCT-2003 (revised)
DT	05-AUG-2003 (first entry)
XX	
DE	HIV-1 JR-FL SOS gp140 amino acid sequence.
XX	
KW	Human immunodeficiency virus; envelope glycoprotein trimeric complex;
KW	HIV; anti-HIV; vaccine; immune response; HIV infection; gp120; gp41;
XX	gp140; furin-recognition sequence.
OS	Human immunodeficiency virus 1.
XX	
PN	WO2003022869-A2.
XX	
PD	20-MAR-2003.
XX	
PF	06-SEP-2002; 2002WO-US028331.
XX	
PR	06-SEP-2001; 2001US-0317764P.
PR	06-SEP-2001; 2001US-0317775P.
PR	06-SEP-2001; 2001US-0317909P.
PR	06-SEP-2001; 2001US-0317910P.
PR	05-APR-2002; 2002US-0370264P.
PR	05-APR-2002; 2002US-0370410P.
XX	
PA	(PROG-) PROGENICS PHARM INC.
PA	(CORR ) CORNELL RES FOUND INC.
XX	
PI	Moore JP, Binley JM, Lu M, Olson WC, Schulke N, Gardner J;
PI	Maddon PJ, Sanders R;
XX	
DR	WPI; 2003-371744/35.
DR	N-PSDB; ACC79597.
XX	
PT	Novel stable HIV-1 pre-fusion envelope glycoprotein trimeric complex in
PT	which each monomeric unit of the complex comprises HIV-1 gp120 and HIV-1
PT	gp41, useful for eliciting immune response in subject against HIV-1.
XX	Example; Fig 13B; 316pp; English.

Search completed: February 25, 2004, 14:16:31  
Job time : 54.8474 secs

```

XX The present invention describes a stable HIV-1 pre-fusion envelope
CC glycoprotein trimeric complex (I), where (i) each monomeric unit of (I)
CC comprises HIV-1 gp120 and HIV-1 gp41, (ii) the gp41 has one or more
CC mutations in its N-terminal helix, and (iii) the gp120 and gp41 are bound
CC to each other by at least one disulfide bond between a cysteine residue
CC introduced into the gp120 and a cysteine residue introduced into the
CC gp41. Also described: (1) a composition (II) comprising a particle and
CC (I) operably affixed to it; (2) a vaccine (III) which comprises a
CC therapeutically or prophylactically effective amount of (I) or (II); and
CC (3) producing (II) by contacting a particle with a stable HIV-1 pre-
CC fusion envelope glycoprotein trimeric complex under conditions permitting
CC the complex to become operable affixed to the particle, or by contacting
CC a particle having an agent which binds to a stable HIV-1 pre-fusion
CC envelope glycoprotein trimeric complex under conditions permitting the
CC complex to bind to the agent, and so permitting the complex to become
CC operably affixed to the particle. (I) has anti-HIV activity. (I) or (II)
CC can be used for eliciting an immune response in a subject against HIV-1
CC or an HIV-1 infected cell. The present sequence represents HIV-1 JR-FL
CC SOS gp140, which is used in an example from the present invention.
CC (Updated on 23-OCT-2003 to standardise OS field)
XX
SQ Sequence 643 AA;
Query Match 86.0%; Score 2862; DB 6; Length 643;
Best Local Similarity 84.5%; Pred. No. 4.2e-138;
Matches 545; Conservative 25; Mismatches 43; Indels 32; Gaps 6;
QY 3 VEKLVTVYVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDNPQEVILENVT 62
DB 1 VEKLVTVYVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDNPQEVILENVT 60
QY 63 ENFMNKNVNNVEOHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNAWTKSNWKMEDR 122
DB 61 EHFNMKNVNNVEOHEDIISLWQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTMR 119
QY 123 GEIKNCSFKVGAG-----KLINCNTSVITQACPKVSPF 155
DB 120 GEIKNCSFNITTSIRDEQKEYALFYKLDVDPDNNNTSYRLISCDSVITQACPKISFE 179
QY 156 PIPHYCAPAGFAILKNDCKKFNKGSTNVSTVQCTHGIRPVVSTQLLNGSLAEEGVV 215
DB 180 PIPHYCAPAGFAILKNDCKTENGKGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVV 239
QY 216 IRSEFTDNAKTIIVQLKESVEINCTRPNNTRKSTIGRPAFYATGDIIGIRQAHCN 275
DB 240 IRSDNETNNAKTIIVQLKESVEINCTRPNNTRKSIHIGRPAFYTTGTEIGIRQAHCN 299
QY 276 ISGKKNNTLKQIVTKLQAFGNKTIIVFKOSGGDPEIVMHSFNCGEFFYCNSTOLFNS 335
DB 300 ISRAKNWDTLQIVIKLREQFENKTIIVFNHSSGGDPEIVMHSFNCGEFFYCNSTOLFNS 359
QY 336 TW-NTTIGPNTNG-TITLPCRILQINRWQEVGKAMYAPPPIRGQIRCCSSNITGLLTRD 393
DB 360 TWNNTEGSNNTGNTITLPCRILQINRWQEVGKAMYAPPPIRGQIRCCSSNITGLLTRD 419
QY 394 GQKEISNTTETIPRGGGMDNRNRSSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVLG 453
DB 420 GGIN-ENGTEIFRPGGMDNRNRSSELYKYKVVKIEPLGVAPTKCKRRVVQREKRAVGIG 478
QY 454 AMFLGFLGAGSTWGAFTSLTVQAROLLSSIVQQQNNLLRAIEAQHLLQLTWGIGK-L 512
DB 479 AVFLGFLGAGSTWGAASMTLTVQARLLLSGI VQQQNNLLRAIEAQRMQLQTLTWGIKQL 538
QY 513 QARVLAVERYLKDQQLLGIWGCSSGKLTCTTAVPWNASWSNKSLLDQIWNNTMTWMEWEREID 572
DB 539 QARVLAVERYLGDQQLLGIWGCSSGKLICTTAVPWNASWSNKSLLDRINWNTMTWMEWEREID 598
QY 573 NYTNLIYTLIEESQNOQEKQELLEDKVASLWNNFDTISKLWY 617
DB 599 NYTSEIYTLIEESQNOQEKQELLEDKVASLWNNFDTITNLWY 643
```

Result No.	Score	Query Match	Length	DB	ID	Description
1	2942.5	88.4	850	2	US-08-448-603A-28	Sequence 28, Appl
2	2942.5	88.4	850	3	US-09-134-075-28	Sequence 28, Appl
3	2942.5	88.4	850	4	US-09-492-739-28	Sequence 28, Appl
4	2880.5	86.6	855	3	US-07-956-483-14	Sequence 14, Appl
5	2876.5	86.5	855	1	US-08-022-835-6	Sequence 6, Appl
6	2876.5	86.5	855	1	US-08-388-809-6	Sequence 6, Appl
7	2876.5	86.5	855	3	US-08-647-714-6	Sequence 6, Appl
8	2851.5	85.7	856	3	US-09-124-900-9	Sequence 9, Appl
9	2851.5	85.7	863	3	US-08-463-210-11	Sequence 11, Appl
10	2851.5	85.7	863	4	US-08-463-028-11	Sequence 11, Appl
11	2848.5	85.6	861	1	US-08-127-499A-14	Sequence 14, Appl
12	2848.5	85.6	861	3	US-08-482-847-14	Sequence 14, Appl
13	2848.5	85.6	861	3	US-07-956-483-10	Sequence 10, Appl
14	2848.5	85.6	861	3	US-08-472-240A-7	Sequence 7, Appl
15	2848.5	85.6	861	3	US-08-472-240A-7	Sequence 7, Appl
16	2848.5	85.6	861	4	US-08-617-441-103	Sequence 103, Appl
17	2848.5	85.6	855	3	US-07-956-483-15	Sequence 15, Appl
18	2844	85.5	857	2	US-08-448-603A-30	Sequence 30, Appl
19	2844	85.5	857	3	US-09-134-075-30	Sequence 30, Appl
20	2844	85.5	857	4	US-09-492-739-30	Sequence 30, Appl
21	2842.5	85.4	880	3	US-08-788-815-7	Sequence 7, Appl
22	2842.5	85.4	880	2	US-09-157-963-7	Sequence 7, Appl
23	2840.5	85.4	861	3	US-07-956-483-16	Sequence 16, Appl
24	2839	85.3	865	3	US-07-956-483-13	Sequence 13, Appl
25	2835	85.2	836	1	US-08-375-510-2	Sequence 2, Appl
26	2835	85.2	836	2	US-08-487-657-2	Sequence 2, Appl
27	2835	85.2	854	4	US-09-309-572-23	Sequence 23, Appl

Db 28 SAAEKLWTVVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIQLN 87  
Qy 61 VTENFNMMKNNMVEQMHEDIISLWDSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 120  
Db 88 VTENFNMMKNNMVEQMHEDIISLWDSLKPCVKLTPLCVTLNCTDLKNAATNTSSWGKM 147  
Qy 121 DRGEIKNCSFKVGAG-----KLINCNTSVITQACPVS 153  
Db 148 ERGEIKNCSFNVTTSIRDKMNEYALFYKLDVVPIDNDNTSYRLISCNTSVITQACPVS 207  
Qy 154 FEPPIHYCAPAGFAILKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAEEG 213  
Db 208 FEPPIHYCAPAGFAILKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAEE 267  
Qy 214 VVIRSENFDTNACTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAPYATGDIIGDIRQAH 273  
Db 268 VVIRSENFDTNACTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAPYATGDIIGDIRQAH 327  
Qy 274 CNISGEKNWNTLKQIVTKLQAOFGNKTIIVFKOSSGGDPEIVMHSFNCGGEFFYCNSTQLF 333  
Db 328 CNLSSTKNWNTLKQIVTKLREHF-NKTIIVFNHSSGGDPEIVMHSFNCGGEFFYCNSTPLF 386  
Qy 334 NS-----TWNTTIGPNNNGTITLPCRQIKQIINRMQEVGKAMYAPPPIRGQIRCSSNITG 387  
Db 387 NSTWNTYTWNTEGSDNTGRNITLQCRQIKQIINRMQEVGKAMYAPPPIRGQIRCSSNITG 446  
Qy 388 LLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKKRVVQREK 447  
Db 447 LLLTRDGGN--NSETEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKKRVVQREK 504  
Qy 448 RAVTLGAMFLGFLGAAGSTMGARSRLTIVQAROLLSGIVQQNNLLRAIEAQHLLQLTV 507  
Db 505 RAVGIGAVFLGFLGAAGSTMGARSRLTIVQAROLLSGIVQQNNLLRAIEAQHLLQLTV 564  
Qy 508 WGIK-LQARVLAVERYLKDQQLLGWCSGKLICTTAVPMNASWSNKSLDQIWNNTWME 566  
Db 565 WGIKQLOARVLAVERYLKDQQLLGWCSGKLICTTAVPMNASWSNKSLDKIWDNNTWME 624  
Qy 567 WEREIDNTNLIYTLIEESQOQEKNEQELLELDKWASLWNNWFDISKWLWYIK 619  
Db 625 WEREIDNTSLIYSLIEESQOQEKNEQELLELDKWASLWNNWFDITKWLWYIK 677

## RESULT 2

US-09-134-075-28  
; Sequence 28, Application US/09134075  
; Patent No. 6042836  
; GENERAL INFORMATION:  
; APPLICANT: Berman, Phillip W.  
; APPLICANT: Nakamura, Gerald R.  
; TITLE OF INVENTION: HIV Envelope Polypeptides  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP  
; STREET: 3 Embarcadero Center  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/134,075  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/448,603  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haliday, Emily

; REGISTRATION NUMBER: 38903  
; REFERENCE/DOCKET NUMBER: 14918-704  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-393-2000  
; TELEFAX: 415-393-2286  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 850 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-134-075-28  
  
Query Match 88.4%; Score 2942.5; DB 3; Length 850;  
Best Local Similarity 85.3%; Pred. No. 8.1e-234;  
Matches 557; Conservative 27; Mismatches 32; Indels 37; Gaps 5;  
  
Qy 1 SAVEKLWTVVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIQLN 60  
Db 28 SAAEKLWTVVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIQLN 87  
Qy 61 VTENFNMMKNNMVEQMHEDIISLWDSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 120  
Db 88 VTENFNMMKNNMVEQMHEDIISLWDSLKPCVKLTPLCVTLNCTDLKNAATNTSSWGKM 147  
Qy 121 DRGEIKNCSFKVGAG-----KLINCNTSVITQACPVS 153  
Db 148 ERGEIKNCSFNVTTSIRDKMNEYALFYKLDVVPIDNDNTSYRLISCNTSVITQACPVS 207  
Qy 154 FEPPIHYCAPAGFAILKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAEEG 213  
Db 208 FEPPIHYCAPAGFAILKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAEE 267  
Qy 214 VVIRSENFDTNACTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAPYATGDIIGDIRQAH 273  
Db 268 VVIRSENFDTNACTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAPYATGDIIGDIRQAH 327  
Qy 274 CNISGEKNWNTLKQIVTKLQAOFGNKTIIVFKOSSGGDPEIVMHSFNCGGEFFYCNSTQLF 333  
Db 328 CNLSSTKNWNTLKQIVTKLREHF-NKTIIVFNHSSGGDPEIVMHSFNCGGEFFYCNSTPLF 386  
Qy 334 NS-----TWNTTIGPNNNGTITLPCRQIKQIINRMQEVGKAMYAPPPIRGQIRCSSNITG 387  
Db 387 NSTWNTYTWNTEGSDNTGRNITLQCRQIKQIINRMQEVGKAMYAPPPIRGQIRCSSNITG 446  
Qy 388 LLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKKRVVQREK 447  
Db 447 LLLTRDGGN--NSETEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKKRVVQREK 504  
Qy 448 RAVTLGAMFLGFLGAAGSTMGARSRLTIVQAROLLSGIVQQNNLLRAIEAQHLLQLTV 507  
Db 505 RAVGIGAVFLGFLGAAGSTMGARSRLTIVQAROLLSGIVQQNNLLRAIEAQHLLQLTV 564  
Qy 508 WGIK-LQARVLAVERYLKDQQLLGWCSGKLICTTAVPMNASWSNKSLDQIWNNTWME 566  
Db 565 WGIKQLOARVLAVERYLKDQQLLGWCSGKLICTTAVPMNASWSNKSLDKIWDNNTWME 624  
Qy 567 WEREIDNTNLIYTLIEESQOQEKNEQELLELDKWASLWNNWFDISKWLWYIK 619  
Db 625 WEREIDNTSLIYSLIEESQOQEKNEQELLELDKWASLWNNWFDITKWLWYIK 677

## RESULT 3

US-09-492-739-28  
; Sequence 28, Application US/09492739  
; Patent No. 6331404  
; GENERAL INFORMATION:  
; APPLICANT: Berman, Phillip W.  
; APPLICANT: Nakamura, Gerald R.  
; TITLE OF INVENTION: HIV Envelope Polypeptides  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:

```

, FILING DATE: 31-DEC-1992
, CLASSIFICATION: 435
, PRIOR APPLICATION NO: 92/19742
, APPLICATION NUMBER: 92/19742
, FILING DATE: 12-NOV-1992
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: FR 91 05392
, FILING DATE: 02-MAY-1991
, ATTORNEY/AGENT INFORMATION:
, NAME: Crane-Feuzy, Sharon E
, REGISTRATION NUMBER: 36,113
, REFERENCE/DOCKET NUMBER: 017753-005
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (703) 836-6620
, TELEFAX: (703) 836-2021
, INFORMATION FOR SEQ ID NO: 14:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 855 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: DNA (genomic)
US-07-956-483-14

Query Match          86.6%; Score 2880.5; DB 3; Length 855;
Best Local Similarity 83.3%; Pred. No. le-228;
Matches 547; Conservative 31; Mismatches 40; Indels 39; Gaps 7

Qy      1  SAVEKLVWTVVYGVVPVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVILN 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      28  SAAEQLVWTVVYGVVPVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVVLN 87
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy      61  VTENFNFNKNNWVQEMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLK-----NATNTKSS 115
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      88  VTENFNFNKNNWVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTLRNDTSNATNTTSS 147
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy      116  NWKEMDRGETKNCSEFKVAG-----KLNCNNTSVITQACP 150
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      148  NRGMEGEGMTNCSFNSTISRSIKVQKEYALFYKLDVVVPIDNTSYTLINCNTSVITQACP 207
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

QY 151 KVSFEPIPIHYCAPAGFAILLKNDKKFNGSGPCTNVSTVQCTHIGIRPVVSTOLLNGSLA 210
Db 208 KVSFEPIPIHYCA-RWFAILLNCNKKFNGSGPCTNVSTVQCTHIGIRPVVSTHLLNGSLA 266
QY 211 EECWVIRSENFTDNAKTIIVQLKESVEINCTRENNNTRKSIITIGPGRAFATGDIIGDIR 270
Db 267 EEEVLRSENFTDNAKTIIVQLKEAVEINCTRENNNTRSIHIGPGRAFATGDIIGDIR 326
QY 271 QAHNCISGEKNWNTLQIVTKLQAQFGNKTIIVFKQSSGGDPEIVMHSFNCGGEFFVCNST 330
Db 327 QAHNCISRAKNWNTLQIVTKLQDQFENKTIIFNRSSGGDPEIVMHSFNCGGEFFVCNST 386
QY 331 QLPNSTWNTTIGPNNTG--TITLPCRIKQIINRWQSVGKAMYAPPPIRGQIRCSSNITGL 388
Db 387 QLPFSSTWNGTEGNSNTGNDTITLPCRIKEIINRWQSVGKAMYAPPPIKGQVCCSSNITGL 446
QY 389 LLTRDGGKEISNT---TEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKERVVQ 444
Db 447 LLTRDGGNSKNGSKNENTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKERVVQ 506
QY 445 REKRAV-TLGAMFLGFLGAAGSTWGAARSULTVQARQLLSGIVQOQNLLRAIEAQOHL 503
Db 507 REKRAVGTIGAMFLGFLGAAGSTWGAATSMTLTQARLLSGIVQOQNLLRAIEAQOHL 566
QY 504 QLTWGIK-LOARVLAVERYLKDQQLLGIWCGSKLICTTAVPWNASWNSKSLDOIWNM 562
Db 567 QLTWGIKQLQARVLAVERYLRDQQLLGIWCGSKLICTTVPWNASWNSKSLDKIWNM 626
QY 563 TWMWEREIDNYTLNLYTLIEESONQOEKNEQELLELDKWSLWNNWFDISKWLWYIK 619
Db 627 TWMWEREIDNYTLNLYTLIEESONQOEKNEQELLELDKWSLWNNWFDITNWLWYIK 683

```

RESULT 5

```

US-08-022-835-6
; Sequence 6, Application US/08022835
; Patent No. 5420030
; GENERAL INFORMATION:
; APPLICANT: Reitz Jr., Marvin S.
; APPLICANT: Franchini, Genoveffa
; APPLICANT: Markham, Phillip D.
; APPLICANT: Gallo, Robert C.
; APPLICANT: Lori, Franco C.
; APPLICANT: Popovic, Mikulas
; APPLICANT: Gartner, Suzanne
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Eleventh Floor, 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/022,835
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/599,491
; FILING DATE: 17-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-022-835-6

Query Match      86.5%; Score 2876.5; DB 1; Length 855;
Best Local Similarity 82.7%; Pred. No. 2.2e-228;
Matches 542; Conservative 33; Mismatches 43; Indels 37; Gaps 5;

QY 1 SAVEKLWTVVYGVVPWKAEATTTLCASADAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
Db 28 NAEKELWTVVYGVVPWKAEATTTLCASDRKAYDTEVHNWATHACVPTDPNPQEVLEN 87
QY 61 VTENFNWKNMVEQWHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNAATNTKSNWKEM 120
Db 88 VTENFNWKNMVEQWHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNAATNTKSNWKEM 147
QY 121 DR-----GEIKCSPKVGAG-----KLINCNSTSVITQ 147
Db 148 SRGWGGGEMKCNFNITNIRGKVQKEVALFYKLDIADIDNNSNNRRLISCNTSVITQ 207
QY 148 ACPKVSFEPIPIHYCAPAGFAILLKNDKKFNGSGPCTNVSTVQCTHIGIRPVVSTOLLNG 207
Db 208 ACPKVSFEPIPIHYCAPAGFAILLKNDKKFNGSGPCTNVSTVQCTHIGIRPVVSTOLLNG 267
QY 208 SLAEGWVIRSNFTDNAKTIIVQLKESVEINCTRENNNTRKSIITIGPGRAFATGDIIG 267
Db 268 SLAEEVWVIRSANFADNAKIIIVQLNESVEINCTRENNNTRKSIHIGPGRAFATGDIIG 327
QY 268 DIRQAHNCISGEKNWNTLQIVTKLQAQFGNKTIIVFKQSSGGDPEIVMHSFNCGGEFFYC 327
Db 328 DIRQAHNCISRAKNWNTLQIVTKLQDQFENKTIIVFKSHSSGGDPEIVTHSFNCGGEFFYC 387
QY 328 NSTQLFNSTWNTTIGPNNT--NGTITLPCRIKQIINRWQSVGKAMYAPPPIRGQIRCSSNI 385
Db 388 NSTQLFNSTWNTTIGPNNT--NGTITLPCRIKQIINRWQSVGKAMYAPPPIRGQIRCSSNI 447
QY 386 TGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKERVVQ 445
Db 448 TGLLLTRDGGPE-DNKTEVFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKERVVQ 506
QY 446 EKRAVTLGAMFLGFLGAAGSTWGAARSULTVQARQLLSGIVQOQNLLRAIEAQOHL 505
Db 507 EKRAVGIGAVFLGFLGAAGSTWGAARSULTVQARQLLSGIVQOQNLLRAIEAQOHL 566
QY 506 TWGVIK-LOARVLAVERYLKDQQLLGIWCGSKLICTTAVPWNASWNSKSLDOIWNMTW 564
Db 567 TWGVIKQLQARVLAVERYLRDQQLLGIWCGSKLICTTAVPWNASWNSKSLDKIWNMTW 626
QY 565 MWEEREIDNYTLNLYTLIEESONQOEKNEQELLELDKWSLWNNWFDISKWLWYIK 619
Db 627 IEWDREINNYTLNLYTLIEESONQOEKNEQELLELDKWSLWNNWFDITNWLWYIK 681

RESULT 6
US-08-388-809-6
; Sequence 6, Application US/08388809
; Patent No. 5576000
; GENERAL INFORMATION:
; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,
; APPLICANT: GENOVEFFA, MARKHAM, PHILIP D. GALLO, ROBERT
; APPLICANT: C. LORI, FRANCO C.; POPOVIC, MIKULAS; AND
; APPLICANT: GARTNER, SUZANNE
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK, 3.5"  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/388,809  
FILING DATE: 15-FEB-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/022,835  
FILING DATE: 25-FEB-1993  
APPLICATION NUMBER: US 07/599,491  
FILING DATE: 17-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: LESLIE A. SERUNIAN  
REGISTRATION NUMBER: 35,353  
REFERENCE/DOCKET NUMBER: 2026-4092U52  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 855 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-388-809-6

Query Match 86.5%; Score 2876.5; DB 1; Length 855;  
Best Local Similarity 82.7%; Pred. No. 2.2e-228;  
Matches 542; Conservative 33; Mismatches 43; Indels 37; Gaps 5;

QY 1 SAVEKLVTVVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTDPNPQEVLEN 60  
DB 28 NAEKLVTVVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTDPNPQEVLEN 87  
QY 61 VTENFNWKNMVMVEQMHEDIISLWDSQKPCVKLTPLCVTLHCTNLKATNTKSSNWKEM 120  
DB 88 VTENFNWKNMVMVEQMHEDIISLWDSQKPCVKLTPLCVTLHCTNLKATNTKSSNWKEM 147  
QY 121 DR-----GEIKNSFKVGAG-----KLIINCTSVITQ 147  
DB 148 SRGMVGGGEMKNSFNITTNIRGKQKEYALFYKLDIAPIDNNNNRRLISCTSVITQ 207  
QY 148 ACPKVSFEPIPIHYPAGAPAILKNDKFKNGSGPCTNVSTVCTHGRPVVSTQLLNG 207  
DB 208 ACPKVSFEPIPIHYPAGAPAILKNDKFKNGSGPCTNVSTVCTHGRPVVSTQLLNG 267  
QY 208 SLAEEGVIRSENFTDNAKTIIVQLKESVEINCTRNNTNRKSIITIGPGRAFATGDIIG 267  
DB 268 SLAEEGVIRSANFADNAKVIIVQLNESVEINCTRNNTNRKSIHIGPGRAFATGDIIG 327  
QY 268 DIRQAHNCISGEKWNNTLQIVTKLQAOFGNKTIVFKQSGGDPETVWHSFNGCGEFFYC 327  
DB 328 DIRQAHNCISRAKWNNTLQIVTKLQAOFGNKTIVFKQSGGDPETVWHSFNGCGEFFYC 387  
QY 328 NSTQLENSTWNNTIGPNNNT--NGTITLPCRIKQIINRWQEVGKAMVAPPTRGQIRCSSNI 385  
DB 388 NSTQLENSTWNNTIGPNNNT--NGTITLPCRIKQIINRWQEVGKAMVAPPTRGQIRCSSNI 447  
QY 386 TGLLTRDGGKEISNTTEIFRPGGGDMRDNRSSELYKYKVKIEPLGVAPTAKRRVQR 445  
DB 448 TGLLTRDGGPE--DNKTEVFRPGGGDMRDNRSSELYKYKVKIEPLGVAPTAKRRVQR 506  
QY 446 EKRAVTLGAMFLGFLGAGSTMGARSLTITVQARQLLSGIVQOQNLLRAIEAQHLLQL 505

Db 507 EKRAVGIGAVFLGFLGAGSTMGAAAMTUTVQARLLLSGIVQOQNLLRAIEAQHLLQL 566  
QY 506 TVWGIK-LQARVLAVERYLKDQQLLGIWCSGKLICTTAVPWNASHSNKSLDQIWNMTW 564  
Db 567 TVWGIKQLQARVLAVERYLRDQQLLGIWCSGKLICTTAVPWNASHSNKSLKNDMTW 626  
QY 565 MEWERIDNYTNLIYTLIESNQOQKNEQELLELDKWSLWNNWFDISKWLWIK 619  
Db 627 IEWDREINNYTSIYSLIESNQOQKNEQELLELDKWSLWNNWFDITKWLWIK 681

RESULT 7  
US-08-647-714-6  
; Sequence 6, Application US/08647714  
; Patent No. 5869313  
; GENERAL INFORMATION:  
; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,  
; APPLICANT: GENOVEFFA; MARKHAM, PHILIP D. GALLO, ROBERT  
; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND  
; APPLICANT: GARTNER, SUZANNE  
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK, 3.5"  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/647,714  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/388,809  
; FILING DATE: 15-FEB-1995  
; APPLICATION NUMBER: US 08/022,835  
; FILING DATE: 25-FEB-1993  
; APPLICATION NUMBER: US 07/599,491  
; FILING DATE: 17-OCT-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LESLIE A. SERUNIAN  
; REGISTRATION NUMBER: 35,353  
; REFERENCE/DOCKET NUMBER: 2026-4092U52  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 855 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-647-714-6

Query Match 86.5%; Score 2876.5; DB 2; Length 855;  
Best Local Similarity 82.7%; Pred. No. 2.2e-228;  
Matches 542; Conservative 33; Mismatches 43; Indels 37; Gaps 5;

QY 1 SAVEKLVTVVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTDPNPQEVLEN 60  
DB 28 NAEKLVTVVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTDPNPQEVLEN 87  
QY 61 VTENFNWKNMVMVEQMHEDIISLWDSQKPCVKLTPLCVTLHCTNLKATNTKSSNWKEM 120  
DB 88 VTENFNWKNMVMVEQMHEDIISLWDSQKPCVKLTPLCVTLHCTNLKATNTKSSNWKEM 147

```

QY 121 DR-----GBIKNSFKVGAG-----KLINCNTSVITQ 147
Db 148 SRGWGGGEMKNSFNITNIRKVKQKEYALFYKLDIAPIDNNSNNRYELISCTSVITQ 207
QY 148 ACPKVSFEPIPIHYCAPAGFAILKCNDDKFNKSGPCTNVSTVQCTHGIRPVVSTQLLNG 207
Db 208 ACPKVSFEPIPIHYCAPAGFAILKCKDKKFNKSGPCTNVSTVQCTHGIRPVVSTQLLNG 267
QY 208 SLAEEGVIRSENFTNAKTIIVOLKESVEINCTRNNTNRKSIITGPGRAFATGDIIG 267
Db 268 SLAEEVWIRSANFADNAKVIIVOLNESVEINCTRNNTNRKSIHIGPGRAFYTTEIGIG 327
QY 268 DIRQAHNTSGEKNNTLKQIVTKLQAOFGNKTIIVFKQSGGDPETVHMSFNCGGFFYC 327
Db 328 DIRQAHNLSRAKNWDTLKNIVILKREQGNTIIVFKHSGGDPETVHMSFNCGGFFYC 387
QY 328 NSTQLFNSTWNTTIGPNNT--NGTITLPCRIKQIINRWQEVGKAMYAPPPIRGQIRCSSNI 385
Db 388 NSTQLFNSTWNTTESNNTVENNTITLPCRIKQIINRWQEVGRAMVAPPPIRGQIRCSSNI 447
QY 386 TGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTAKARRVOR 445
Db 448 TGLLLTRDGGPE--DNKTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTAKARRVOR 506
QY 446 EKRAVTLGAMFLGFLGAAGSTMARSILTLVQARQLLSGIVQOQNLLRAIEAQOHLLOL 505
Db 507 EKRAVGIGAVFLGFLGAGSTMGRAAWTLTVQARLLLSGIVQOQNLLRAIEAQOHLLOL 566
QY 506 TVMGIK--LQARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASNSKSLDOJWNMTW 564
Db 567 TVMGIKQARVLAVERYLRDQQLLGWCGSGKLICTTAVPWNASNSKSLNKINDWMTW 626
QY 565 MEWEREIDNTNLYTLIEESQOQEKNEQELLELDKWSLWNNFDISKWLWYIK 619
Db 627 IEWDREINNTSYIISLIEESQOQEKNEQELLELDKWSLWNNFDTTKWLWYIK 681

```

RESULT 8

```

US-09-124-900-9
; Sequence 9, Application US/09124900
; Patent No. 6268484
; GENERAL INFORMATION:
; APPLICANT: KATINGER, Hermann
; APPLICANT: BUCHACHER, Andrea
; APPLICANT: ERNST, Wolfgang
; APPLICANT: BALLAUN, Claudia
; APPLICANT: PURTSCHER, Martin
; APPLICANT: TRKOLA, Alexandra
; APPLICANT: PREDL, Renate
; APPLICANT: SCHMATZ, Christine
; APPLICANT: KLIMA, Annelies
; APPLICANT: STEINDL, Franz
; APPLICANT: MUSTER, Thomas
; TITLE OF INVENTION: HIV-Vaccines
; FILE REFERENCE: 1939-112P
; CURRENT APPLICATION NUMBER: US/09/124,900
; CURRENT FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: PCT/EP95/01481
; PRIOR FILING DATE: 1995-04-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 856
; TYPE: PR1
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (79)..(184)
; NAME/KEY: PEPTIDE
; LOCATION: (326)..(400)
US-09-124-900-9
Query Match 85.7%; Score 2851; DB 3; Length 856;

```

```

Best Local Similarity 83.4%; Pred. No. 2.8e-226;
Matches 548; Conservative 32; Mismatches 37; Indels 40; Gaps 10;

QY 1 SAVEKLWTVYVYGVPMKEATTTLCASDAKAYDTEVHNWATHACVPTDPPQETVLEN 60
Db 29 SATEKLWTVYVYGVPMKEATTTLCASDAKAYDTEVHNWATHACVPTDPPQEVVLVN 88
QY 61 VTENFNMMQNNVQOMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWE- 119
Db 89 VTENFNMMQNNVQOMHEDIISLWQSLKPCVKLTPLCVLSLACTDLKNDTNTNSSGRMI 148
QY 120 MDGEIKNCSFKVGA---GK-----LINCNTSVITQCPKV 152
Db 149 MEKGEIKNCSFNISIRGKQKEYAFFYKLDIIPIDNDTTSVTLTSCNTSVITQCPKV 208
QY 153 SPEPIPIHYCAPAGFAILKCNDDKFNKSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAE 212
Db 209 SPEPIPIHYCAPAGFAILKCNKNTFGTGPCNTVSTVQCTHGIRPVVSTQLLNGSLAE 268
QY 213 GVVISSEFTDNAKTIIVOLKESVEINCTRNNTNRKSIIT--GPGRAFATGDIICDIR 270
Db 269 EVVISSEFTDNAKTIIVOLNQSVEINCTRNNTNRKSIIRIQGPGRAFVTIGK--IGNMR 327
QY 271 QAHCNISGEKNNTLKQIVTKLQAOFG--NKTIVFKQSGGDPETVHMSFNCGGFFYCNS 329
Db 328 QAHCNISRAKNWTLKQIDSKLREQFGNNKTIIFKQSGGDPETVHMSFNCGGFFYCNS 387
QY 330 TOLFSTWNT--TIGPNNTNG--TITLPCRIKQIINRWQEVGKAMYAPPPIRGQIRCSS 383
Db 388 TOLFSTWNTSTWSTKGSNTTSDTITLPCRIKQIINRWQEVGKAMYAPPISGQIRCSS 447
QY 384 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTAKARRVV 443
Db 448 NITGLLLTRDGGNS--NNESEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTAKARRVV 506
QY 444 QREKRAVTLGAMFLGFLGAAGSTMARSILTLVQARQLLSGIVQOQNLLRAIEAQOHLLOL 503
Db 507 QREKRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQOQNLLRAIEAQOHLLOL 566
QY 504 QLTVMGIK--LQARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASNSKSLDOJWNMM 562
Db 567 QLTVMGIKQARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASNSKSLDOJWNMM 626
QY 563 TWMEREIDNTNLYTLIEESQOQEKNEQELLELDKWSLWNNFDISKWLWYIK 619
Db 627 TWMEREIDNTNLYTLIEESQOQEKNEQELLELDKWSLWNNFDTTKWLWYIK 683

```

RESULT 9

```

US-08-463-210-11
; Sequence 11, Application US/08463210
; Patent No. 6001977
; GENERAL INFORMATION:
; APPLICANT: CHANG, Nancy T.
; APPLICANT: GALLO, Robert C.
; APPLICANT: WONG-STAAAL, Flossie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,210
; FILING DATE: 05-JUN-1995

```



CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/693,866  
FILING DATE: 23-JAN-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/659,339  
FILING DATE: 10-OCT-1984  
ATTORNEY/AGENT INFORMATION:  
NAME: Serunian, Lealie A.  
REGISTRATION NUMBER: 35,353  
REFERENCE/DOCKET NUMBER: 2026-4193US2  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 863 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: HTLV-III  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..863  
OTHER INFORMATION: /note="env protein of HTLV-III"  
US-08-463-210-11  
Query Match 85.7%; Score 2851; DB 3; Length 863;  
Best Local Similarity 83.4%; Pred. No. 2.8e-226;  
Matches 548; Conservative 32; Mismatches 37; Indels 40; Gaps 10;  
Qy 1 SAVEKLWTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVVLN 60  
Db 36 SATEKLWTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVVLN 95  
Qy 61 VTENFNWKNMVEQMHEDIISLWDSLKPCVKLTPLCVTLKTNLKNATNTKSSNWK- 119  
Db 96 VTENFNWKNMVEQMHEDIISLWDSLKPCVKLTPLCVTLKTNLKNATNTKSSNWK- 155  
Qy 120 MDRGEIKNCSFKVGA---GK-----LINCNTSVITQACPKV 152  
Db 156 MERGEIKNCSFNISTIRGKVEYAFYKLDIIPDNDTTSYTLTSCNTSVITQACPKV 215  
Qy 153 SFEPPIHYCAPAGFALLKNDKFKSGPCTNVTQCTHGRPVVSTQLLNGSLAE 212  
Db 216 SFEPPIHYCAPAGFALLKNNKTFNGTGPCTNVTQCTHGRPVVSTQLLNGSLAE 275  
Qy 213 GVVRSENFDTNARTIIVQLKESVEINCTRPNNTRKSI--GPGRAFVATGDIIGDIR 270  
Db 276 EVVIRSANFTDNARTIIVQLNQSVEINCTRPNNTRKSIHQGPGRAFVTIGK-IGNWR 334  
Qy 271 QAHCNISGEKNNTLKOIVTKLAQFG-NKTIIVFKOSSGDPPIVHMSFNCGGEFFYCN 329  
Db 335 QAHCNISRAKNNTLQKIDSKLREQFGNKNKTIIFKQSSGDPPIVTHSFNCGGEFFYCN 394  
Qy 330 TQLFNSTWNN---TIGPNNTNG--TITLPCRKQIINRWQEVGKAMYPPIRGQIRCSS 383  
Db 395 TQLFNSTWNTSTWSTGNNTEGSDTITLPCRKQIINRWQEVGKAMYPPIRGQIRCSS 454  
Qy 384 NITGLLLTRDGGKEISNTTIFRPGGDMRDNRSLEYKVKVIRPLGVAPTKARRVV 443  
Db 455 NITGLLLTRDGGNS--NNESEIFRPGGDMRDNRSLEYKVKVIRPLGVAPTKARRVV 513  
Qy 444 QREKRAVTLGAMFLGFLGAGSTMGARSFLTVOARQLLSGIVQQQNNLLRAIEAQOHL 503  
Db 514 QREKRAVGIGALFLGFLGAGSTMGASMTLTVOARQLLSGIVQQQNNLLRAIEAQOHL 573  
Qy 504 QLTWVGK-LQARVLAVERYLKQQLLGIWGCSSGKLICTTAVPWNASWNSKSLDQIWNM 562  
Db 574 QLTWVGKILQARVLAVERYLKQQLLGIWGCSSGKLICTTAVPWNASWNSKSLDQIWNM 633

Qy 563 TWMWEREIDNNTLYITLIESQOQEKNEQELLELDKWSLWNNWFDISKWLWYIK 619  
Db 634 TWMWDRREINNTSLIHSLIESQOQEKNEQELLELDKWSLWNNWFINITNLWLWYIK 690  
RESULT 10  
US-08-463-028-11  
Sequence 11, Application US/08463028  
Patent No. 6610476  
GENERAL INFORMATION:  
APPLICANT: CHANG, Nancy T.  
APPLICANT: GALLO, Robert C.  
TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,028  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/693,866  
FILING DATE: 23-JAN-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/659,339  
FILING DATE: 10-OCT-1984  
ATTORNEY/AGENT INFORMATION:  
NAME: Serunian, Lealie A.  
REGISTRATION NUMBER: 35,353  
REFERENCE/DOCKET NUMBER: 2026-4193US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 863 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: HTLV-III  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..863  
OTHER INFORMATION: /note="env protein of HTLV-III"  
US-08-463-028-11  
Query Match 85.7%; Score 2851; DB 4; Length 863;  
Best Local Similarity 83.4%; Pred. No. 2.8e-226;  
Matches 548; Conservative 32; Mismatches 37; Indels 40; Gaps 10;  
Qy 1 SAVEKLWTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVVLN 60  
Db 36 SATEKLWTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVVLN 95  
Qy 61 VTENFNWKNMVEQMHEDIISLWDSLKPCVKLTPLCVTLKTNLKNATNTKSSNWK- 119  
Db 96 VTENFNWKNMVEQMHEDIISLWDSLKPCVKLTPLCVTLKTNLKNATNTNTSSGRMI 155

```

QY 120 MDRGEIKNCSFKVGA---GK-----LINCNTSVITQACPKV 152
Db 156 MEKGEIKNCSFNISTIRGKVQKEYAFFYKLDIIPIDNDTTSYTLSCNTSVITQACPKV 215
QY 153 SPEPIPIHYCAPAGFAILKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQLLINGSLABE 212
Db 216 SPEPIPIHYCAPAGFAILKNNKTFGTGCTNVSTVQCTHGIRPVVSTQLLINGSLABE 275
QY 213 GUVIRSENFNDNAKTIIVQLKESVEINCTRPNNNTRKSITI--GPGRAFATGDIIGDIR 270
Db 276 EVVIRSANFTDANKTIIVQLNQSVSEINCTRPNNNTRKSIRIORGPGRAFVTIGK-IGNMR 334
QY 271 QAHCNISGKKNWNTLQKIVTKLQAQFG-NKTIIVFKQSSGGDPBIVMHSFNCGGEPFYCNS 329
Db 335 QAHCNISRAKWNNTLQIDSKLREQFGNKNKTIIFKQSSGGDPBIVTHSFNCGGEPFYCNS 394
QY 330 TQLFNSFTWNN---TIGPNNTNG--TITLPCRIOIINRWQEVGKAMYAPPPIRGQIRCSS 383
Db 395 TQLFNSFTWNSWTSGSNNTGSDTITLPCRIOIINRWQEVGKAMYAPPISGQIRCSS 454
QY 384 NITGLLTRDGGKEISNTTEIFRPGGDMRDNWRSSELYKYKVVKIEPLGVAPTAKKRVV 443
Db 455 NITGLLTRDGGNS--NNESEIFRPGGDMRDNWRSSELYKYKVVKIEPLGVAPTAKKRVV 513
QY 444 QREKRAVTLGAMFLGFLGAAGSTMGARSLLTIVQARQLLSGIVQOQNLLRAIEAQOHL 503
Db 514 QREKRAVGIGALFLGFLGAAGSTMGASMTLTVQARQLLSGIVQOQNLLRAIEAQOHL 573
QY 504 QLTVMGK-LQARVLAVERYLKQDLGIGWCSGKLICTTAVPWNASWSNKSJDQIWN 562
Db 574 QLTVMGKIQLOARILAVERYLKQDLGIGWCSGKLICTTAVPWNASWSNKSLEQIWN 633
QY 563 TWMWEAREIDNYNLTYLIESQNOQKNEQBELLEDKWSLWNNFIDISKWLWYIK 619
Db 634 TWMWEAREIDNYNLTYLIESQNOQKNEQBELLEDKWSLWNNFIDISKWLWYIK 690

```

RESULT 11

```

US-08-127-499A-14
; Sequence 14, Application US/08127499A
; Patent No. 5510264
; GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; APPLICANT: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
; TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 25,768
; REFERENCE/DOCKET NUMBER: 51916/102/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEFAX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
US-08-127-499A-14

Query Match      85.6%; Score 2848.5; DB 1; Length 861;
Best Local Similarity 82.8%; Pred. No. 4.5e-226;
Matches 548; Conservative 30; Mismatches 39; Indels 45; Gaps 10;

QY 1 SAVEKLWTVVYGVPMKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQETVLEN 60
Db 29 SATEKLWTVVYGVPMKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVVLVN 88
QY 61 VTENFNMMKNVVEQMHEDIISLWDQSLKPCVKLPLCVTLHCTNLKNAATNTKSSNKE- 119
Db 89 VTENFNMMKNVVEQMHEDIISLWDQSLKPCVKLPLCVTLHCTNLKNAATNTSSNTSS 148
QY 120 -----MDRGEIKNCSFKVGA---GK-----LINCNTSVITQ 147
Db 149 SGEEMMEKEIKNCSFNISTIRGKVQKEYAFFYKLDIIPIDNDTTSYTLSCNTSVITQ 208
QY 148 ACPKVSFPIPIHYCAPAGFAILKNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQLLNG 207
Db 209 ACPKVSFPIPIHYCAPAGFAILKNNKTFGTGCTNVSTVQCTHGIRPVVSTQLLNG 268
QY 208 SLAEEGVVIRSENFNDNAKTIIVQLKESVEINCTRPNNNTRKSITI--GPGRAFATGDI 265
Db 269 SLAEEGVVIRSENFNDNAKTIIVQLNQSVSEINCTRPNNNTRKSIRIORGPGRAFVTIGK- 327
QY 266 IGDIRQAHCNISGKKNWNTLQKIVTKLQAQFG-NKTIIVFKQSSGGDPBIVMHSFNCGGEP 324
Db 328 IGMRQAHCNISRAKWNNTLQIDSKLREQFGNKNKTIIFKQSSGGDPBIVTHSFNCGGEP 387
QY 325 FYCNSLTQFNSFTWNN---TIGPNNTNG--TITLPCRIOIINRWQEVGKAMYAPPPIRGQ 378
Db 388 FYCNSLTQFNSFTWNSWTSGSNNTGSDTITLPCRIOIINRWQEVGKAMYAPPISGQ 447
QY 379 IRCSSNITGLLTRDGGKEISNTTEIFRPGGDMRDNWRSSELYKYKVVKIEPLGVAPTAK 438
Db 448 IRCSSNITGLLTRDGGNN--NNGSEIFRPGGDMRDNWRSSELYKYKVVKIEPLGVAPTAK 506
QY 439 KRVVQREKRAVTLGAMFLGFLGAAGSTMGARSLLTIVQARQLLSGIVQOQNLLRAIEA 498
Db 507 KRVVQREKRAVGIGALFLGFLGAAGSTMGARSMTLTVQARQLLSGIVQOQNLLRAIEA 566
QY 499 QOHLQLTVMGK-LQARVLAVERYLKQDLGIGWCSGKLICTTAVPWNASWSNKSLEQ 557
Db 567 QOHLQLTVMGKIQLOARILAVERYLKQDLGIGWCSGKLICTTAVPWNASWSNKSLEQ 626
QY 558 IWNNTWMEAREIDNYNLTYLIESQNOQKNEQBELLEDKWSLWNNFIDISKWLWY 617
Db 627 IWNNTWMEAREIDNYNLTYLIESQNOQKNEQBELLEDKWSLWNNFIDISKWLWY 686
QY 618 IK 619
Db 687 IK 688

RESULT 12
US-08-482-847-14
; Sequence 14, Application US/08482847
; Patent No. 5556757
; GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; APPLICANT: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR
; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington

```

```
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/482,847
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/127,499
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 51916/104/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-482-847-14

Query Match      85.6%; Score 2848.5; DB 1; Length 861;
Best Local Similarity 82.8%; Pred. No. 4.5e-226;
Matches 548; Conservative 30; Mismatches 39; Indels 45; Gaps 10;

QY      1 SAVEKLWTVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
DB      29 SATEKLWTVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 88
QY      61 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAATNTKSSNKE- 119
DB      89 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAATNTKSSNKE- 148
QY      120 -----MDRGEIKNCSPKVA---GK-----LINCNTSVITQ 147
DB      149 SGEEMMEKEIKNCSPNISTIRGKVQKEYAFYKLDIIPIDNDTTSYLTSCNTSVITQ 208

; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/482,847
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/127,499
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 51916/104/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-482-847-14

Query Match      85.6%; Score 2848.5; DB 3; Length 861;
Best Local Similarity 82.8%; Pred. No. 4.5e-226;
Matches 548; Conservative 30; Mismatches 39; Indels 45; Gaps 10;

QY      1 SAVEKLWTVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
DB      29 SATEKLWTVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 88
QY      61 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAATNTKSSNKE- 119
DB      89 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAATNTKSSNKE- 148
QY      120 -----MDRGEIKNCSPKVA---GK-----LINCNTSVITQ 147
DB      149 SGEEMMEKEIKNCSPNISTIRGKVQKEYAFYKLDIIPIDNDTTSYLTSCNTSVITQ 208

; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/482,847
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/127,499
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 51916/104/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-482-847-14

Query Match      85.6%; Score 2848.5; DB 3; Length 861;
Best Local Similarity 82.8%; Pred. No. 4.5e-226;
Matches 548; Conservative 30; Mismatches 39; Indels 45; Gaps 10;

QY      1 SAVEKLWTVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
DB      29 SATEKLWTVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 88
QY      61 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAATNTKSSNKE- 119
DB      89 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAATNTKSSNKE- 148
QY      120 -----MDRGEIKNCSPKVA---GK-----LINCNTSVITQ 147
DB      149 SGEEMMEKEIKNCSPNISTIRGKVQKEYAFYKLDIIPIDNDTTSYLTSCNTSVITQ 208
```



```

; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,240A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,483
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..831
; US-08-472-240A-7

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:18:41 ; Search time 32.2957 Seconds  
(without alignments)  
4047.104 Million cell updates/sec

Title: US-09-891-609A-4  
Perfect score: 3327  
Sequence: 1 SAVEKLWVTYYGVVWKEA.....DKWASLWNWFDISKWLWYIK 619

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3327	100.0	619	9	US-09-891-609-4
2	3327	100.0	646	9	US-09-891-609-2
3	3275	98.4	842	14	US-10-190-435-2
4	3275	98.4	842	14	US-10-241-009-2
5	3275	98.4	842	14	US-10-190-434B-2
6	3275	98.4	842	14	US-10-190-305A-2
7	3275	98.4	847	9	US-09-476-242-2
8	2862	86.0	643	14	US-10-032-162-13
9	2848.5	85.6	861	14	US-10-026-741-103
10	2834	85.2	856	9	US-09-476-242-1
11	2819	84.7	856	14	US-10-196-515-11
12	2813.5	84.6	858	9	US-09-938-406-1
13	2797.5	84.1	726	14	US-10-196-515-3
14	2793.5	84.0	759	14	US-10-196-515-12
15	2771	83.3	579	14	US-10-032-162-15

16	2758	82.9	625	14	US-10-032-162-17	Sequence 17, Appl
17	2642.5	79.4	860	14	US-10-190-435-6	Sequence 6, Appl1
18	2642.5	79.4	860	14	US-10-241-009-6	Sequence 6, Appl1
19	2642.5	79.4	860	14	US-10-190-434B-6	Sequence 6, Appl1
20	2642.5	79.4	860	14	US-10-190-305A-6	Sequence 6, Appl1
21	2637.5	79.3	853	13	US-10-003-035-33	Sequence 33, Appl
22	2637.5	79.3	853	13	US-10-286-332A-33	Sequence 33, Appl
23	2637.5	79.3	853	15	US-10-280-915-33	Sequence 33, Appl
24	2588	77.8	870	14	US-10-190-435-147	Sequence 147, App
25	2585	77.7	867	14	US-10-190-435-3	Sequence 3, Appl1
26	2585	77.7	867	14	US-10-190-435-126	Sequence 126, App
27	2585	77.7	867	14	US-10-241-009-3	Sequence 3, Appl1
28	2585	77.7	867	14	US-10-190-434B-3	Sequence 3, Appl1
29	2585	77.7	867	14	US-10-190-305A-3	Sequence 3, Appl1
30	2582.5	77.6	865	14	US-10-190-435-140	Sequence 140, App
31	2579	77.5	858	14	US-10-190-435-150	Sequence 150, App
32	2573.5	77.4	855	14	US-10-190-435-144	Sequence 144, App
33	2570	77.2	862	14	US-10-190-435-141	Sequence 141, App
34	2566.5	77.1	870	14	US-10-190-435-127	Sequence 127, App
35	2564	77.1	869	14	US-10-241-009-4	Sequence 4, Appl1
36	2564	77.1	869	14	US-10-190-434B-4	Sequence 4, Appl1
37	2564	77.1	869	14	US-10-190-305A-4	Sequence 4, Appl1
38	2564	77.1	869	14	US-10-190-435-139	Sequence 139, App
39	2558	76.9	861	14	US-10-190-435-129	Sequence 129, App
40	2555.5	76.8	845	14	US-10-190-435-130	Sequence 130, App
41	2555.5	76.8	845	14	US-10-190-435-134	Sequence 134, App
42	2554.5	76.8	803	14	US-10-190-435-134	Sequence 134, App
43	2551.5	76.7	803	14	US-10-190-435-135	Sequence 135, App
44	2547.5	76.6	857	14	US-10-190-435-138	Sequence 138, App
45	2539.5	76.3	845	14	US-10-190-435-143	Sequence 143, App

ALIGNMENTS

RESULT 1  
US-09-891-609-4  
; Sequence 4, Application US/09891609  
; Patent No. US20020127238A1  
; GENERAL INFORMATION:  
; APPLICANT: Stamatos, Leonidas  
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR  
; FILE REFERENCE: 2570-1-002N  
; CURRENT APPLICATION NUMBER: US/09/891,609  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/214,608  
; PRIOR FILING DATE: 2000-06-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 619  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-891-609-4

Query Match	100.0%;	Score 3327;	DB 9;	Length 619;
Best Local Similarity	100.0%;	Pred. No. 2.8e-296;		
Matches 619;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	SAVEKLWVTYYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN	60	
Db	1	SAVEKLWVTYYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN	60	
Qy	61	VTFENFNKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM	120	
Db	61	VTFENFNKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM	120	
Qy	121	DRGEIKNCSPKVGAGKGLINCNTSVITQACPKVSEFPIPIHYCAPAGFAILKCNCKKFNKS	180	
Db	121	DRGEIKNCSPKVGAGKGLINCNTSVITQACPKVSEFPIPIHYCAPAGFAILKCNCKKFNKS	180	
Qy	181	GPCTNVSTVQCTHGIRPVWSTQLLNGSLAEGVWVIRSENFDTNAKTIIIVQLKESVIEINC	240	

```
Db 181 GPCNTNSTVQCTHGIRPVVSTOLLNGLSLAEEGWIRSENFNTDANKTIIIVQLKESVEINC 240
Qy 241 TRPNNTKRSITIGPGRAFYATGDIIGDIRQAHNCISGEKWNNTLKQIVTKLQAOFGNKT 300
Db 241 TRPNNTKRSITIGPGRAFYATGDIIGDIRQAHNCISGEKWNNTLKQIVTKLQAOFGNKT 300
Qy 301 IVFKSSGGDPEIVMHSFNCGGFFCYCNSTOLFNSWTNNNTIGPNNNTNGTITILPCRIKQII 360
Db 301 IVFKSSGGDPEIVMHSFNCGGFFCYCNSTOLFNSWTNNNTIGPNNNTNGTITILPCRIKQII 360
Qy 361 NRQEVGKAMYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSEL 420
Db 361 NRQEVGKAMYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSEL 420
Qy 421 YKVKVKIEPLGVAPTAKARRVVOREKRAVTIGAMFLGFLGAGSTMGARSJTLTVQARQ 480
Db 421 YKVKVKIEPLGVAPTAKARRVVOREKRAVTIGAMFLGFLGAGSTMGARSJTLTVQARQ 480
Qy 481 LLSGIVQOQNLLRAIEAOQHLLQLTVWGIKQARVLAVERYLKQOQLLGIWGCCKLIC 540
Db 481 LLSGIVQOQNLLRAIEAOQHLLQLTVWGIKQARVLAVERYLKQOQLLGIWGCCKLIC 540
Qy 541 TTAVPNASWSNKSQDIWNNTMTWEEREIDNYTNLIYTLIEESQOQEKNEQELLELD 600
Db 541 TTAVPNASWSNKSQDIWNNTMTWEEREIDNYTNLIYTLIEESQOQEKNEQELLELD 600
Qy 601 KWASLWNWFDISKWLWYIK 619
Db 601 KWASLWNWFDISKWLWYIK 619

RESULT 2
US-09-891-609-2
; Sequence 2, Application US/09891609
; Patent No. US20020127238A1
; GENERAL INFORMATION:
; APPLICANT: Statatatos, Leondias
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
; FILE REFERENCE: 2570-1-002N
; CURRENT APPLICATION NUMBER: US/09/891,609
; PRIOR FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-891-609-2

Query Match 100.0%; Score 3327; DB 9; Length 646;
Best Local Similarity 100.0%; Pred. No. 3e-296;
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAVEKLWTVYVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNQEIVLEN 60
Db 28 SAVEKLWTVYVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNQEIVLEN 87
Qy 61 VTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 120
Db 88 VTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 147
Qy 121 DRGEIKNCSFKVAGKGLINCNTSVITQACPKVSFEPIPIHYCAPAGAPAILKCNCKKFNKS 180
Db 148 DRGEIKNCSFKVAGKGLINCNTSVITQACPKVSFEPIPIHYCAPAGAPAILKCNCKKFNKS 207
Qy 181 GPCNTNSTVQCTHGIRPVVSTOLLNGLSLAEEGWIRSENFNTDANKTIIIVOLKESVEINC 240
Db 208 GPCNTNSTVQCTHGIRPVVSTOLLNGLSLAEEGWIRSENFNTDANKTIIIVOLKESVEINC 267
Qy 241 TRPNNTKRSITIGPGRAFYATGDIIGDIRQAHNCISGEKWNNTLKQIVTKLQAOFGNKT 300
```

```
Db 268 TRPNNTKRSITIGPGRAFYATGDIIGDIRQAHNCISGEKWNNTLKQIVTKLQAOFGNKT 327
Qy 301 IVFKSSGGDPEIVMHSFNCGGFFCYCNSTOLFNSWTNNNTIGPNNNTNGTITILPCRIKQII 360
Db 328 IVFKSSGGDPEIVMHSFNCGGFFCYCNSTOLFNSWTNNNTIGPNNNTNGTITILPCRIKQII 387
Qy 361 NRQEVGKAMYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSEL 420
Db 388 NRQEVGKAMYAPPPIRGQIRCSNITGLLLTRDGGKEISNTTTEIFRPGGDMRDNRSEL 447
Qy 421 YKVKVKIEPLGVAPTAKARRVVOREKRAVTIGAMFLGFLGAGSTMGARSJTLTVQARQ 480
Db 448 YKVKVKIEPLGVAPTAKARRVVOREKRAVTIGAMFLGFLGAGSTMGARSJTLTVQARQ 507
Qy 481 LLSGIVQOQNLLRAIEAOQHLLQLTVWGIKQARVLAVERYLKQOQLLGIWGCCKLIC 540
Db 508 LLSGIVQOQNLLRAIEAOQHLLQLTVWGIKQARVLAVERYLKQOQLLGIWGCCKLIC 567
Qy 541 TTAVPNASWSNKSQDIWNNTMTWEEREIDNYTNLIYTLIEESQOQEKNEQELLELD 600
Db 568 TTAVPNASWSNKSQDIWNNTMTWEEREIDNYTNLIYTLIEESQOQEKNEQELLELD 627
Qy 601 KWASLWNWFDISKWLWYIK 619
Db 628 KWASLWNWFDISKWLWYIK 646

RESULT 3
US-10-190-435-2
; Sequence 2, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SF162
US-10-190-435-2

Query Match 98.4%; Score 3275; DB 14; Length 842;
Best Local Similarity 95.2%; Pred. No. 2.6e-291;
Matches 616; Conservative 0; Mismatches 3; Indels 28; Gaps 2;

Qy 1 SAVEKLWTVYVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNQEIVLEN 60
Db 23 SAVEKLWTVYVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNQEIVLEN 82
Qy 61 VTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 120
Db 83 VTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM 142
Qy 121 DRGEIKNCSFKVAGKGLINCNTSVITQACPKVSFEPIPIHYCAPAGAPAILKCNCKKFNKS 153
Db 143 DRGEIKNCSFKVAGKGLINCNTSVITQACPKVSFEPIPIHYCAPAGAPAILKCNCKKFNKS 202
Qy 154 FEPIPIHYCAPAGAPAILKCNCKKFNKSGPCTNVSTVQCTHGIRPVVSTOLLNGLSLAEEG 213
Db 203 FEPIPIHYCAPAGAPAILKCNCKKFNKSGPCTNVSTVQCTHGIRPVVSTOLLNGLSLAEEG 262
Qy 214 VVIRSENFNTDANKTIIIVOLKESVEINCTRPNNNTKRSITIGPGRAFYATGDIIGDIRQAH 273
```



```
Db 263 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAFYATGDIIGDIRQAH 322
QY 274 CNISGEKWNNTLKOIVTKLQAOFGNKTIVFKQSSGGDPPEIVMHSFNCGGGFFCYCNSTOLF 333
Db 323 CNISGEKWNNTLKOIVTKLQAOFGNKTIVFKQSSGGDPPEIVMHSFNCGGGFFCYCNSTOLF 382
QY 334 NSTWNTNIGPNTNGTITLPCRILKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRD 393
Db 383 NSTWNTNIGPNTNGTITLPCRILKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRD 442
QY 394 GGKEISNTTIEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVTLG 453
Db 443 GGKEISNTTIEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVTLG 502
QY 454 AMFLGFLGAAGSTWGARSLLTVQARQLLSGIVQOQNLLRAIEAQOHLQLTVWGK-L 512
Db 503 AMFLGFLGAAGSTWGARSLLTVQARQLLSGIVQOQNLLRAIEAQOHLQLTVWGK-L 562
QY 513 QARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASWSNKSJDQIWNNTMMEWEREID 572
Db 563 QARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASWSNKSJDQIWNNTMMEWEREID 622
QY 573 NYTNLIYTLIEESQOQEKNEQELLELDKWSLWNNFDISKWLWYIK 619
Db 623 NYTNLIYTLIEESQOQEKNEQELLELDKWSLWNNFDISKWLWYIK 669

RESULT 4
US-10-241-009-2
; Sequence 2, Application US/10241009
; Publication No. US20030170614A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; FILE REFERENCE: 2300-1621.21
; CURRENT APPLICATION NUMBER: US/10/241,009
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SF162
US-10-241-009-2

Query Match 98.4%; Score 3275; DB 14; Length 842;
Best Local Similarity 95.2%; Pred. No. 2.6e-291;
Matches 616; Conservative 0; Mismatches 3; Indels 28; Gaps 2;

QY 1 SAVEKLWTVTVYGVVPWKAEATTLTFCASDAKAYDTEVHNWATHACVPTDPPNQEIVLEN 60
Db 23 SAVEKLWTVTVYGVVPWKAEATTLTFCASDAKAYDTEVHNWATHACVPTDPPNQEIVLEN 82
QY 61 VTENFNWKNMVEQMHEDIISLWDSKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 120
Db 83 VTENFNWKNMVEQMHEDIISLWDSKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 142
QY 121 DRGEIKNCSPKVGAG-----KLI NCNTSVITQACP KVS 153
Db 143 DRGEIKNCSPKVTTSIRNMKQKEYALFYKLDVVPIDNDNTSYKLI NCNTSVITQACP KVS 202
QY 154 FEPPIHYCAPAGFAILKCNKDKKFNKSGGCTNVSTVQCTHGIRPVVSTQLLNGSLAEEG 213
Db 203 FEPPIHYCAPAGFAILKCNKDKKFNKSGGCTNVSTVQCTHGIRPVVSTQLLNGSLAEEG 262
QY 214 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAFYATGDIIGDIRQAH 273
Db 263 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAFYATGDIIGDIRQAH 322
```

```
Db 263 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAFYATGDIIGDIRQAH 322
QY 274 CNISGEKWNNTLKOIVTKLQAOFGNKTIVFKQSSGGDPPEIVMHSFNCGGGFFCYCNSTOLF 333
Db 323 CNISGEKWNNTLKOIVTKLQAOFGNKTIVFKQSSGGDPPEIVMHSFNCGGGFFCYCNSTOLF 382
QY 334 NSTWNTNIGPNTNGTITLPCRILKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRD 393
Db 383 NSTWNTNIGPNTNGTITLPCRILKQIINRWQEVGKAMYPPIRGQIRCSNITGLLLTRD 442
QY 394 GGKEISNTTIEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVTLG 453
Db 443 GGKEISNTTIEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVTLG 502
QY 454 AMFLGFLGAAGSTWGARSLLTVQARQLLSGIVQOQNLLRAIEAQOHLQLTVWGK-L 512
Db 503 AMFLGFLGAAGSTWGARSLLTVQARQLLSGIVQOQNLLRAIEAQOHLQLTVWGK-L 562
QY 513 QARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASWSNKSJDQIWNNTMMEWEREID 572
Db 563 QARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASWSNKSJDQIWNNTMMEWEREID 622
QY 573 NYTNLIYTLIEESQOQEKNEQELLELDKWSLWNNFDISKWLWYIK 619
Db 623 NYTNLIYTLIEESQOQEKNEQELLELDKWSLWNNFDISKWLWYIK 669

RESULT 5
US-10-190-434B-2
; Sequence 2, Application US/10190434B
; Publication No. US20030194800A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; FILE REFERENCE: 2300-1621.20
; CURRENT APPLICATION NUMBER: US/10/190,434B
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SF162
US-10-190-434B-2

Query Match 98.4%; Score 3275; DB 14; Length 842;
Best Local Similarity 95.2%; Pred. No. 2.6e-291;
Matches 616; Conservative 0; Mismatches 3; Indels 28; Gaps 2;

QY 1 SAVEKLWTVTVYGVVPWKAEATTLTFCASDAKAYDTEVHNWATHACVPTDPPNQEIVLEN 60
Db 23 SAVEKLWTVTVYGVVPWKAEATTLTFCASDAKAYDTEVHNWATHACVPTDPPNQEIVLEN 82
QY 61 VTENFNWKNMVEQMHEDIISLWDSKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 120
Db 83 VTENFNWKNMVEQMHEDIISLWDSKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 142
QY 121 DRGEIKNCSPKVGAG-----KLI NCNTSVITQACP KVS 153
Db 143 DRGEIKNCSPKVTTSIRNMKQKEYALFYKLDVVPIDNDNTSYKLI NCNTSVITQACP KVS 202
QY 154 FEPPIHYCAPAGFAILKCNKDKKFNKSGGCTNVSTVQCTHGIRPVVSTQLLNGSLAEEG 213
Db 203 FEPPIHYCAPAGFAILKCNKDKKFNKSGGCTNVSTVQCTHGIRPVVSTQLLNGSLAEEG 262
QY 214 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAFYATGDIIGDIRQAH 273
Db 263 VVIRSEFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAFYATGDIIGDIRQAH 322
```

QY 274 CNISGEKWNNTLKQIVTKLQAQFGNKTIVFKOSSGGDPEIVMHSFNCGGFFFCYNSTQLP 333  
DB 323 CNISGEKWNNTLKQIVTKLQAQFGNKTIVFKOSSGGDPEIVMHSFNCGGFFFCYNSTQLP 382  
QY 334 NSTWNNTTGNNTNGTITLPCRICKQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRD 393  
DB 383 NSTWNNTTGNNTNGTITLPCRICKQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRD 442  
QY 394 GKKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVVQREKRAVTLG 453  
DB 443 GKKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVVQREKRAVTLG 502  
QY 454 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQQNNLLRAIEAQOHLQLTVWGK-L 512  
DB 503 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQQNNLLRAIEAQOHLQLTVWGK-L 562  
QY 513 QARVLAVERYLKDOQLLGIWGCSSGKLICTTAVPWNASWSNKSLLDQIWNNTWMEWEREID 572  
DB 563 QARVLAVERYLKDOQLLGIWGCSSGKLICTTAVPWNASWSNKSLLDQIWNNTWMEWEREID 622  
QY 573 NYTNLIYTLIEESQNOQEKNEQELLELDKWSLWNNFDISKWLWYIK 619  
DB 623 NYTNLIYTLIEESQNOQEKNEQELLELDKWSLWNNFDISKWLWYIK 669

RESULT 6  
US-10-190-305A-2  
; Sequence 2, Application US/10190305A  
; Publication No. US20030198621A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MESEDE, Jan  
; APPLICANT: BARNETT, Susan  
; APPLICANT: LIAN, Ying  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR  
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 2302-18702 / 18702.002  
; CURRENT APPLICATION NUMBER: US/10/190,305A  
; CURRENT FILING DATE: 2002-07-05  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 842  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SF162  
US-10-190-305A-2

Query Match 98.4%; Score 3275; DB 14; Length 842;  
Best Local Similarity 95.2%; Pred. No. 2.6e-291;  
Matches 616; Conservative 0; Mismatches 3; Indels 28; Gaps 2;  
QY 1 SAVEKLVWTVVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60  
DB 23 SAVEKLVWTVVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 82  
QY 61 VTENFNWKNMVEQMHEDIISLWDSQSKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 120  
DB 83 VTENFNWKNMVEQMHEDIISLWDSQSKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 142  
QY 121 DRGEIKNCSPKVGAG-----KLIINCTSVITQACPVS 153  
DB 143 DRGEIKNCSPKVTTSIRNKQKEYALFYKLDVVPIDNDNTSYKLIINCTSVITQACPVS 202  
QY 154 FEPIPIHYCAPAGFAILKCNCKKFNCGSPCTNVTSTVQCTHGIRPVVSTQLLNGSLAEEG 213  
DB 203 FEPIPIHYCAPAGFAILKCNCKKFNCGSPCTNVTSTVQCTHGIRPVVSTQLLNGSLAEEG 262  
QY 214 VVIRSENFTDNAKTIIIVQLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAH 273  
DB 263 VVIRSENFTDNAKTIIIVQLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAH 322

QY 274 CNISGEKWNNTLKQIVTKLQAQFGNKTIVFKOSSGGDPEIVMHSFNCGGFFFCYNSTQLP 333  
DB 323 CNISGEKWNNTLKQIVTKLQAQFGNKTIVFKOSSGGDPEIVMHSFNCGGFFFCYNSTQLP 382  
QY 334 NSTWNNTTGNNTNGTITLPCRICKQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRD 393  
DB 383 NSTWNNTTGNNTNGTITLPCRICKQIINRWQEVGKAMYPPIRGQIRCSSNITGLLLTRD 442  
QY 394 GKKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVVQREKRAVTLG 453  
DB 443 GKKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVVQREKRAVTLG 502  
QY 454 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQQNNLLRAIEAQOHLQLTVWGK-L 512  
DB 503 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQQNNLLRAIEAQOHLQLTVWGK-L 562  
QY 513 QARVLAVERYLKDOQLLGIWGCSSGKLICTTAVPWNASWSNKSLLDQIWNNTWMEWEREID 572  
DB 563 QARVLAVERYLKDOQLLGIWGCSSGKLICTTAVPWNASWSNKSLLDQIWNNTWMEWEREID 622  
QY 573 NYTNLIYTLIEESQNOQEKNEQELLELDKWSLWNNFDISKWLWYIK 619  
DB 623 NYTNLIYTLIEESQNOQEKNEQELLELDKWSLWNNFDISKWLWYIK 669

RESULT 7  
US-09-476-242-2  
; Sequence 2, Application US/09476242  
; Patent No. US20020146683A1  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: HARTOG, Karin  
; APPLICANT: MARTIN, Eric  
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
; FILE REFERENCE: 1605.002  
; CURRENT APPLICATION NUMBER: US/09/476,242  
; CURRENT FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 847  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-476-242-2

Query Match 98.4%; Score 3275; DB 9; Length 847;  
Best Local Similarity 95.2%; Pred. No. 2.6e-291;  
Matches 616; Conservative 0; Mismatches 3; Indels 28; Gaps 2;

QY 1 SAVEKLVWTVVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60  
DB 28 SAVEKLVWTVVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 87  
QY 61 VTENFNWKNMVEQMHEDIISLWDSQSKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 120  
DB 88 VTENFNWKNMVEQMHEDIISLWDSQSKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEM 147  
QY 121 DRGEIKNCSPKVGAG-----KLIINCTSVITQACPVS 153  
DB 148 DRGEIKNCSPKVTTSIRNKQKEYALFYKLDVVPIDNDNTSYKLIINCTSVITQACPVS 207  
QY 154 FEPIPIHYCAPAGFAILKCNCKKFNCGSPCTNVTSTVQCTHGIRPVVSTQLLNGSLAEEG 213  
DB 208 FEPIPIHYCAPAGFAILKCNCKKFNCGSPCTNVTSTVQCTHGIRPVVSTQLLNGSLAEEG 267  
QY 214 VVIRSENFTDNAKTIIIVQLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAH 273  
DB 268 VVIRSENFTDNAKTIIIVQLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAH 327  
QY 274 CNISGEKWNNTLKQIVTKLQAQFGNKTIVFKOSSGGDPEIVMHSFNCGGFFFCYNSTQLP 333  
DB 328 CNISGEKWNNTLKQIVTKLQAQFGNKTIVFKOSSGGDPEIVMHSFNCGGFFFCYNSTQLP 387

QY 334 NSTWNTIGNNTNGTITLPCRKIQIINRWQEVGKAMYAPPPIRGQIRCCSSNITGLLLTRD 393  
Db 388 NSTWNTIGNNTNGTITLPCRKIQIINRWQEVGKAMYAPPPIRGQIRCCSSNITGLLLTRD 447  
QY 394 GGKEISNTTTEIFRPGGDMRNRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVTLG 453  
Db 448 GGKEISNTTTEIFRPGGDMRNRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVTLG 507  
QY 454 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQOQNLLRAEAQOHLQLTVWGK-L 512  
Db 508 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQOQNLLRAEAQOHLQLTVWGK-L 567  
QY 513 QARVLAVERYLKDQQLLGIWGCSSKLICTTAVPWNASWSNKSJDQIWNNTWMEWEREID 572  
Db 568 QARVLAVERYLKDQQLLGIWGCSSKLICTTAVPWNASWSNKSJDQIWNNTWMEWEREID 627  
QY 573 NYTNLYTTLIESQOQEKNEQLLELDKWSLWNNFDDISKWLYIK 619  
Db 628 NYTNLYTTLIESQOQEKNEQLLELDKWSLWNNFDDISKWLYIK 674

RESULT 8  
US-10-032-162-13  
; Sequence 13, Application US/10032162  
; Publication No. US2003005283A1  
; GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADDON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331a2  
; CURRENT APPLICATION NUMBER: US/10/032,162  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 13  
; LENGTH: 643  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-032-162-13

Query Match 86.0%; Score 2862; DB 14; Length 643;  
Best Local Similarity 84.5%; Pred. No. 1.4e-253;  
Matches 545; Conservative 25; Mismatches 43; Indels 32; Gaps 6;

QY 3 VEKLWTVVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVILENT 62  
Db 1 VEKLWTVVYGVVPWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVILENT 60  
QY 63 ENFMNKNVQOQHEHDIISLWQSLKPCVKLTPLCVTLHCTNLKATNTKSNWKEMDR 122  
Db 61 EHFNMKNVQOQHEHDIISLWQSLKPCVKLTPLCVTLHCTNLKATNTKSNWKEMDR 119  
QY 123 GETKNCSEFKVGAG-----KLINCNTSVITQACPKVSPE 155  
Db 120 GETKNCSEFKVGAG-----KLINCNTSVITQACPKVSPE 179  
QY 156 PIPHYCAPAGPAILKCNKDKFNGSGPCTNVSTVQCTHGRPVVSTQLLINGSLEEGVV 215  
Db 180 PIPHYCAPAGPAILKCNKDKFNGSGPCTNVSTVQCTHGRPVVSTQLLINGSLEEGVV 239  
QY 216 IRSENFDTNAKTIIVOLKESVEINCTRPNNNTKSTIGPGAFYATGDIIGDIAHNCN 275  
Db 240 IRSENFDTNAKTIIVOLKESVEINCTRPNNNTKSTIGPGAFYATGDIIGDIAHNCN 299  
QY 276 ISGEKKNNTLKQIVTKLQAFQGNKTVFKQSSGDDPEIIVHWSFNCGEFFYCNSTOLFNS 335  
Db 300 ISRAKWNDTLKQIVTKLQAFQGNKTVFKQSSGDDPEIIVHWSFNCGEFFYCNSTOLFNS 359

QY 336 TW-NNTIGNNTNG-TITLPCRKIQIINRWQEVGKAMYAPPPIRGQIRCCSSNITGLLLTRD 393  
Db 360 TWNNNTGNSNTGNTITLPCRKIQIINRWQEVGKAMYAPPPIRGQIRCCSSNITGLLLTRD 419  
QY 394 GGKEISNTTTEIFRPGGDMRNRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVTLG 453  
Db 420 GGIN-ENGTEIFRPGGDMRNRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIG 478  
QY 454 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQOQNLLRAEAQOHLQLTVWGK-L 512  
Db 479 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNLLRAEAQOHLQLTVWGK-L 538  
QY 513 QARVLAVERYLKDQQLLGIWGCSSKLICTTAVPWNASWSNKSJDQIWNNTWMEWEREID 572  
Db 539 QARVLAVERYLKDQQLLGIWGCSSKLICTTAVPWNASWSNKSJDQIWNNTWMEWEREID 598  
QY 573 NYTNLYTTLIESQOQEKNEQLLELDKWSLWNNFDDISKWLYIK 617  
Db 599 NYTSEIYTLIESQOQEKNEQLLELDKWSLWNNFDDITNMLWY 643

RESULT 9  
US-10-026-741-103  
; Sequence 103, Application US/10026741  
; Publication No. US20030049604A1  
; GENERAL INFORMATION:  
; APPLICANT: CHARNEAU, PIERRE  
; CLAVEL, FRANCOISE  
; BORMAN, ANDREW  
; QUILLIENT, CAROLINE  
; GUETARD, DENISE  
; MONTAGNIER, LUC  
; DONJON DE SAINT-MARTIN, JACQUELINE  
; COHEN, JACQUES  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; Dunner, L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/026,741  
; FILING DATE: 27-Dec-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/817,441  
; FILING DATE: 31-AUG-1998  
; APPLICATION NUMBER: PCT/FR 95/01391  
; FILING DATE: 20-OCT-1995  
; APPLICATION NUMBER: FR 9412554  
; FILING DATE: 20-OCT-1994  
; APPLICATION NUMBER: FR 9502526  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03260.6005-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 103:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 861 amino acids  
; TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 103:
US-10-026-741-103

Query Match      85.6%; Score 2848.5; DB 14; Length 861;
Best Local Similarity 82.8%; Pred. No. 3.8e-252;
Matches 548; Conservative 30; Mismatches 39; Indels 45; Gaps 10;

QY 1 SAVEKLWTVTVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPPQEVVLN 60
DB 29 SATEKLWTVTVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPPQEVVLN 88
QY 61 VTENFNMMKNNMVEQMHEDIISLWDSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWE- 119
DB 89 VTENFNMMKNDMVEQMHEDIISLWDSLKPCVKLTPLCVSLKCTDLGNATNTSSNWS 148
QY 120 -----MDRGEIKNCSFKVGA--GK-----LINCNTSVITQ 147
DB 149 SGEMMEKEIKNCSFNISTSIRGKQVEYAFYKLDIIPIDNDTTSYLTSCNTSVITQ 208
QY 148 ACPKVSFEPIPIHYCAPAGAILKCNKPKNGSGPCTNVSTVQCTHGIRPVVSTQLLLNG 207
DB 209 ACPKVSFEPIPIHYCAPAGAILKCNKPTNGTGPCTNVSTVQCTHGIRPVVSTQLLLNG 268
QY 208 SLAEEGVIRSENFDTNAKTIIVOLKESVEINCTRPNNTRKSIITI--GPGRAFATYD 265
DB 269 SLAEEVIRSENFDTNAKTIIVOLKESVEINCTRPNNTRKSIIRIQPGRAFVITGK- 327
QY 266 IGIROAHNCNISGEKWNNTLKQIVTKLQAFG-NKTIIVFKQSSGGDPEIVMHSFNCGGEF 324
DB 328 IGWRQAHCNISRAKNWNTLKQIASKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEF 387
QY 325 FYCNSQLFNSWN-----TIGPNNTNG--TITLPCRIKQIINRWQEVGKAMYAPPIRQ 378
DB 388 FYCNSQLFNSWNSTWSTEGSNNTGSDTITLPCRIKQIINRWQEVGKAMYAPPIRQ 447
QY 379 IRCSSNITGLLTRDGGKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAK 438
DB 448 IRCSSNITGLLTRDGGNN--NNESEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAK 506
QY 439 KRRVOREKRAVTLGAMFLGFLGAAGSTMARSILTLVQARQLLSGIVQOQNLLRAIEA 498
DB 507 KRRVOREKRAVGIGALFLGFLGAAGSTMARSILTLVQARQLLSGIVQOQNLLRAIEA 566
QY 499 QHLLQLTVMGIK-LQARVLAVERYLKDQQLLGIWCGSKLICTTAVPWNASWSNKSLEQ 557
DB 567 QHLLQLTVMGIKQARILAVERYLKDQQLLGIWCGSKLICTTAVPWNASWSNKSLEQ 626
QY 558 IWNMTWMEWERIDNTNLIYTLIEESQOEKNEQELLELDKWSLWNPFDISKWLY 617
DB 627 IWNMTWMEWDREINNTYLSLHLSIESQOEKNEQELLELDKWSLWNPFDISKWLY 686
QY 618 IK 619
DB 687 IK 688

RESULT 10
US-09-476-242-1
; Sequence 1, Application US/09476242
; Patent No. US2002014683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-1

Query Match      85.2%; Score 2834; DB 9; Length 856;
Best Local Similarity 82.6%; Pred. No. 8e-251;
Matches 543; Conservative 33; Mismatches 41; Indels 40; Gaps 9;

QY 1 SAVEKLWTVTVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPPQEVVLN 60
DB 29 SATEKLWTVTVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPPQEVVLN 88
QY 61 VTENFNMMKNNMVEQMHEDIISLWDSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWE- 119
DB 89 VTENFNMMKNDMVEQMHEDIISLWDSLKPCVKLTPLCVSLKCTDLKNDTNTSSGRMI 148
QY 120 MDRGEIKNCSFKVGA-----KLINCNTSVITQACPKV 152
DB 149 MEKGEIKNCSFNISTSIRGKQVEYAFYKLDIIPIDNDTTSYLTSCNTSVITQACPKV 208
QY 153 SPEPIPIHYCAPAGAILKCNKPKNGSGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAE 212
DB 209 SPEPIPIHYCAPAGAILKCNKPTNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAE 268
QY 213 GVVISSENFDTNAKTIIVOLKESVEINCTRPNNTRKSIITI--GPGRAFATYDGIIDIR 270
DB 269 EVVISSENFDTNAKTIIVOLKESVEINCTRPNNTRKSIIRIQPGRAFVITGK-IONMR 327
QY 271 QAHCNISGEKWNNTLKQIVTKLQAFG-NKTIIVFKQSSGGDPEIVMHSFNCGGEFYCN 329
DB 328 QAHCNISRAKNWNTLKQIASKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFYCN 387
QY 330 TQLFNSWN-----TIGPNNTNG--TITLPCRIKQIINRWQEVGKAMYAPPIRQIRCSS 383
DB 388 TQLFNSWNSTWSTEGSNNTGSDTITLPCRIKQIINRWQEVGKAMYAPPIRQIRCSS 447
QY 384 NITGLLTRDGGKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVV 443
DB 448 NITGLLTRDGGNS--NNESEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVV 506
QY 444 QREKRAVTLGAMFLGFLGAAGSTMARSILTLVQARQLLSGIVQOQNLLRAIEAQOHL 503
DB 507 QREKRAVGIGALFLGFLGAAGSTMARSILTLVQARQLLSGIVQOQNLLRAIEAQOHL 566
QY 504 QLTWVGIK-LQARVLAVERYLKDQQLLGIWCGSKLICTTAVPWNASWSNKSLEQIWNM 562
DB 567 QLTWVGIKQARILAVERYLKDQQLLGIWCGSKLICTTAVPWNASWSNKSLEQIWNHT 626
QY 563 TWMEWERIDNTNLIYTLIEESQOEKNEQELLELDKWSLWNPFDISKWLYIK 619
DB 627 TWMEWDREINNTYLSLHLSIESQOEKNEQELLELDKWSLWNPFDISKWLYIK 683

RESULT 11
US-10-196-515-11
; Sequence 11, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24

```

```
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

Query Match      84.7%; Score 2819; DB 14; Length 856;
Best Local Similarity 82.5%; Pred. No. 1.9e-249;
Matches 542; Conservative 35; Mismatches 40; Indels 40; Gaps 10;

QY 1 SAVEKLWTVVYGVGVKWEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 NATEKLWTVVYGVGVKWEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 88
QY 61 VTENFNMMKNMVEQMHEDIISLWDSQKPCVKLTPLCVTLHCT--NLKNATNTKSNKE- 119
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
89 VTENFDMKNMVEQMHEDIISLWDSQKPCVKLTPLCVSLKCTDLKNDTNTSSGRMI 148
QY 120 MRGEIKNCSFKVGA---GK-----LINCNTSVITQACPKV 152
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 MEKGEIKNCSFNISTSGRVQKVEYAFYKLDIIPIDNDTTSYSLSCNTSVITQACPKV 208
QY 153 SPEPIPIHYCAPAGFAILKCNKDKKFGSGPCTNVSTVQCTHGIRPVVSTQLLINGSABE 212
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
209 SPEPIPIHYCAPAGFAILKCNKDKKFGSGPCTNVSTVQCTHGIRPVVSTQLLINGSABE 268
QY 213 GVVISRNFDTNAKTIIVQLKESVEINCTRPNNTRKSTI--GPCRAFYATGDIIGDIR 270
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 EVVISRNFDTNAKTIIVQLNVSVEINCTRPNNTRKRIQKPGRAFTVTK-IGNMR 327
QY 271 QAHCNISGEKNWNTLKQIVTKLQAQFG-NKTIIVFKQSSGGDPEIVMHSFNCGSEFFYCN 329
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 QAHCNISRAKNWNTLKQIDSKLREQFGNKTIIFFKQSSGGDPEIVTHSFNCGSEFFYCN 387
QY 330 TQLFNSTWNN---TIGPNTNG--TITLPCRIKQIINRWQEVGKAMYPPIRGQIRCS 383
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 TQLFNSTWNTWSTGSGNTEGSDTITLPCRIKQIINRWQEVGKAMYPPIRGQIRCS 447
QY 384 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKVKVKBPLGVAPTKARRV 443
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 NITGLLLTRDGGNS--NNESEIFRPGGDMRDNRSELYKVKVKBPLGVAPTKARRV 506
QY 444 QREKRAVTLGAMFLGFLGAAGSTMGARSULTVQARQLLSGIVQOQNLLRAIEAQHLL 503
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
507 QREKRAVGIGALFLGFLGAAGSTMGASNTLTVQARQLLSGIVQOQNLLRAIEAQHLL 566
QY 504 QLTWVGIK-LQARVLAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDQIWNM 562
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
567 QLTWVGIKQARILAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDQIWNHT 626
QY 563 TWMEWREIDNTLYTLIEESONQOEKNEQELLELDKWSLWVFDISKWLYIK 619
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
627 TWMEWREINNTLSLHSLIEESONQOEKNEQELLELDKWSLWVFNITNWLWYIK 683

RESULT 12
US-09-938-406-1
; Sequence 1, Application US/09938406
; Patent No. US20020155120A1
; GENERAL INFORMATION:
; APPLICANT: Lowell, George
; APPLICANT: Vancott, Thomas
; APPLICANT: Birx, Deborah
; TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR
; FILE OF INVENTION: INDUCING MUCOSAL IMMUNITY
; FILE REFERENCE: 40646-20002.10
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/938,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/214,701
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/US 97/12253
```

```
; PRIOR FILING DATE: 1997-07-10
; PRIOR APPLICATION NUMBER: US 60/021,687
; PRIOR FILING DATE: 1996-07-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 868
; TYPE: PRT
; ORGANISM: Virus HIV-1
US-09-938-406-1

Query Match      84.6%; Score 2813.5; DB 9; Length 868;
Best Local Similarity 79.3%; Pred. No. 6.2e-249;
Matches 533; Conservative 36; Mismatches 44; Indels 59; Gaps 8;

QY 1 SAVEKLWTVVYGVGVKWEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30 SAANLWTVVYGVGVKWEATTTLCASDAKAYDTEAHNWATHACVPTDPNPQEVLEN 89
QY 61 VTENFNMMKNMVEQMHEDIISLWDSQKPCVKLTPLCVTLHCT--NLKNATNTKSN-- 116
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90 VTENFNMMKNMVEQMHEDIISLWDSQKPCVKLTPLCVTLNCTDLTNTTNTTSLII 149
QY 117 --WKMDRGEIKNCSFKVGA-----KLINCNT 142
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
150 VVMEQRGKGMERNCSFNITTSIRDKVQREYALFYKLDVBPIDDKVNTTNTKYRLIN 209
QY 143 SVITQACPKVSFEPIPIHYCAPAGFAILKCNKDKKFGSGPCTNVSTVQCTHGIRPVVST 202
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
210 SVITQACPKVSFEPIPIHYCTPTGFALLKCNKDKKFGSGPCTNVSTVQCTHGIRPVVST 269
QY 203 LLLNGSLABEAGVVISRNFDTNAKTIIVQLKESVEINCTRPNNTRKSTITIGRAFIAT 262
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 LLLNGSLABEAGVVISRNFDTNAKTIIVQLNVSVEINCTRPNNTRKRVTLGPRVWY 329
QY 263 GDIIGDIRQAHCNISGEKNWNTLKQIVTKLQAQFGNKTIVFKQSSGGDPEIVMHSFNC 332
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
330 GEILGNIRQAHCNISRAQWNNTLQOIATTLREQFGNKTIAFNQSSGGDPEIVMHSFNC 389
QY 323 EFPYCNSTOLFNSWNNTIGPNTNGT-----ITLPCRIKQIINRWQEVGKA 369
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
390 EFPYCNSTOLFNSWNNT-----SNGTWSVTRKQKDTGDIITLPCRIKQIINRWQ 444
QY 370 MYAPPPIRGQIRCSNITGLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKVKVIE 429
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
445 MYALPIKGLIRCSNITGLLTRDGGGE-NQTEIFRPGGDMRDNRSELYKVKVIE 503
QY 430 PLGVAPTKARRVQREKRAV-TLGAMFLGFLGAAGSTMGARSULTVQARQLLSGIVQ 488
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 PLGVAPTKARRVQREKRAVGMFLGFLGAAGSTMGATSMALTQARQLLSGIVQ 563
QY 489 QNNLLRAIEAQHLLQLTWVGIK-LQARVLAVERYLKDQQLLGIWCGSGKLICTTAVPN 547
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
564 QNNLLRAIEAQHLLQLTWVGIKQIARILAVERYLKDQQLLGIWCGSGKLICTTAVPN 623
QY 548 ASWSKSLDQIWNNTWMEWREIDNTLYTLIEESONQOEKNEQELLELDKWSLWV 607
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
624 ASWSKSLDQIWNNTWMEWREIDNTLYTLIEESONQOEKNEQELLELDKWSLWV 683
QY 608 WFDISKWLYIK 619
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
684 WSDITKWLWYIK 695

RESULT 13
US-10-196-515-3
; Sequence 3, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
```

```
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; TITLE OF INVENTION: THERAPEUTICS
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 726
; TYPE: PRF
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-3

Query Match      84.1%; Score 2797.5; DB 14; Length 726;
Best Local Similarity 81.8%; Pred. No. 1.4e-247;
Matches 534; Conservative 37; Mismatches 45; Indels 37; Gaps 9;

QY 1 SAVEKLWVTYYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVLN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 NATEKLWVTYYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVLN 88
QY 61 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAATNTKSSNWE- 119
Db VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVSLKCTDLDKNDTNTSSGGRMI 148
QY 120 MRGEIKNCSFKVGAGK-----LINCSVTITQACPKV 152
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 MEKEIKNCSFNISTSKRGVKVEYAFYKLDIIPIDNDPTSYTLTSCNTSVITQACPKV 208
QY 153 SPEPIPIHYCAPAGFAILKNDKFKNGSGPCTNNVSTVQCTHGIRPVVSTQLLNGSLAE 212
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
209 SPEPIPIHYCAPAGFAILKCNKNTFNGTGTCTNNVSTVQCTHGIRPVVSTQLLNGSLAE 268
QY 213 GVVISRNTDNAKTIIVQLKESVEINCTRPNNNTKRSITI--GCGRAFATGDIIGDIR 270
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 EVVIRSVNFTDNAKTIIVQLNTSVEINCTKPNNTKRIRIHRGPGRAFVTVGK-IGNMR 327
QY 271 QAHCNISGKWNNTLKOIVTKLQAFQ-NKTIIVFKOSSGGDPEIVMHSFNCGGEFFYCS 329
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 QAHCNISRAKWNNTLQIASKLREQFGNNKTIIFKOSSGGDPEIVTHSFNCGGEFFYCKS 387
QY 330 TQLFNSTWNTIGPNNNTG--TITLPCRILQIINRWQEVGKAMYPPIRGQIRCSSNITG 387
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 TQLFNSTW-STKGSNNTEGSDTITLPCRILQIINRWQEVGKAMYPPIRGQIRCSSNITG 446
QY 388 LLLTRDGGKEISNTTTEIFRPGGDMRDNMRSELYKYKVVKIEPLGVAPTAKKRRVVOREK 447
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
447 LLLTRDGGNS--NNSEIFRPGGDMRDNMRSELYKYKVVKIEPLGVAPTAKKRRVVOREK 505
QY 448 RAVTLGAMFLGFLGAAGSTWGARSALTITVOARQLLSGIVQOQNNLLRAIEAQHLLQLTV 507
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
506 RAVGIGALFLGFLGAAGSTWGAASMAITVOARQSLSGIVQOQNNLLRAIEAQHLLQLTV 565
QY 508 WGIK-LQARVLAVERYLKQOQLLGWCSGKLTCTTAVPWNASWSNKSQDQIWNNTWME 566
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
566 WGIKQLQARLAVERYLKQOQLLGWCSGKLTCTTAVPWNASWSNKSLEQIWNNTWME 625
QY 567 WEREIDNTNLIYLTIEESQOQEKNEQELLELDKWSLWNNWFDISKWLWYIK 619
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
626 WDREINNTSLIHSLTIEESQIQOEMNEQELLELDKWSLWNNWFDITNWLWYIK 678

RESULT 14
US-10-196-515-12
; Sequence 12, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celia C.
```

```
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; TITLE OF INVENTION: THERAPEUTICS
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 759
; TYPE: PRF
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-12

Query Match      84.0%; Score 2793.5; DB 14; Length 759;
Best Local Similarity 81.9%; Pred. No. 3.5e-247;
Matches 535; Conservative 36; Mismatches 45; Indels 37; Gaps 9;

QY 1 SAVEKLWVTYYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVLN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 NATEKLWVTYYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVLN 88
QY 61 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAATNTKSSNWE- 119
Db VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVSLKCTDLDKNDTNTSSGGRMI 148
QY 120 MRGEIKNCSFKVGAGK-----LINCSVTITQACPKV 152
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 MEKEIKNCSFNISTSKRGVKVEYAFYKLDIIPIDNDPTSYTLTSCNTSVITQACPKV 208
QY 153 SPEPIPIHYCAPAGFAILKNDKFKNGSGPCTNNVSTVQCTHGIRPVVSTQLLNGSLAE 212
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
209 SPEPIPIHYCAPAGFAILKCNKNTFNGTGTCTNNVSTVQCTHGIRPVVSTQLLNGSLAE 268
QY 213 GVVISRNTDNAKTIIVQLKESVEINCTRPNNNTKRSITI--GCGRAFATGDIIGDIR 270
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 EVVIRSVNFTDNAKTIIVQLNTSVEINCTKPNNTKRIRIHRGPGRAFVTVGK-IGNMR 327
QY 271 QAHCNISGKWNNTLKOIVTKLQAFQ-NKTIIVFKOSSGGDPEIVMHSFNCGGEFFYCS 329
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 QAHCNISRAKWNNTLQIASKLREQFGNNKTIIFKOSSGGDPEIVTHSFNCGGEFFYCKS 387
QY 330 TQLFNSTWNTIGPNNNTG--TITLPCRILQIINRWQEVGKAMYPPIRGQIRCSSNITG 387
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 TQLFNSTW-STKGSNNTEGSDTITLPCRILQIINRWQEVGKAMYPPIRGQIRCSSNITG 446
QY 388 LLLTRDGGKEISNTTTEIFRPGGDMRDNMRSELYKYKVVKIEPLGVAPTAKKRRVVOREK 447
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
447 LLLTRDGGNN--NNSEIFRPGGDMRDNMRSELYKYKVVKIEPLGVAPTAKKRRVVOREK 505
QY 448 RAVTLGAMFLGFLGAAGSTWGARSALTITVOARQLLSGIVQOQNNLLRAIEAQHLLQLTV 507
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
506 RAVGIGALFLGFLGAAGSTWGAASMAITVOARQSLSGIVQOQNNLLRAIEAQHLLQLTV 565
QY 508 WGIK-LQARVLAVERYLKQOQLLGWCSGKLTCTTAVPWNASWSNKSQDQIWNNTWME 566
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
566 WGIKQLQARLAVERYLKQOQLLGWCSGKLTCTTAVPWNASWSNKSLEQIWNNTWME 625
QY 567 WEREIDNTNLIYLTIEESQOQEKNEQELLELDKWSLWNNWFDISKWLWYIK 619
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
626 WDREINNTSLIHSLTIEESQOQEMNEQELLELDKWSLWNNWFDITNWLWYIK 678

RESULT 15
US-10-032-162-15
; Sequence 15, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: LABRANCHE, Celia C.
```

; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADDON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331az  
; CURRENT APPLICATION NUMBER: US/10/032,162  
; PRIOR FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: 09/602,864  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-032-162-15

Query Match 83.3%; Score 2771; DB 14; Length 579;  
Best Local Similarity 85.6%; Pred. No. 2,7e-245;  
Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

QY 3 VEKLVTVYYGVVPWKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVILENVT 62  
Db |||||  
QY 1 VEKLVTVYYGVVPWKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVILENVT 60  
Db |||||  
QY 63 ENFNMMKNNVQEMHEDIISLDQSLKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEMDR 122  
Db |::|  
QY 61 EHFNMKNNVQEMQEDIISLDQSLKPCVKLTPLC----- 96  
Db |::|  
QY 123 GEIKNSFKVGAGKLINCNTSVITQCPKVSPEPIPIHYCAPAGFAILKNDKKFNGSGP 182  
Db |||||  
QY 97 -----GAG-----CDTSVITQCPKISFEPIPIHYCAPAGFAILKNDKTFNGKGP 142  
Db |||||  
QY 183 CTNVSTVQCTHGIRPVVSTOLLNGSLAEGVIRSEFTDNAKTIIVQLKESVEINCTR 242  
Db |||||  
QY 143 CRNVSTVQCTHGIRPVVSTOLLNGSLAEEVIRSDNFTNNAKTIIVQLKESVEINCTR 202  
Db |||||  
QY 243 PNNTRKSIIGRAPHATGDIIGDIRQAHNCISGEKNNTLKQIVTKLOAFGNKTIV 302  
Db |||||  
QY 203 PNNTRKSIHIGRAFYTTGEIIGDIRQAHNCISRAKNNDTLKQIVIKLREQFENKTIV 262  
Db |||||  
QY 303 FKSSGDDPEIVMHSFNCGEFFYCNSTQLFNSTW-NNTIGPNNTNG-TITLPCRICKII 360  
Db |||||  
QY 263 FHHSSGDDPEIVMHSFNCGEFFYCNSTQLFNSTWNNTEGSGNTEGNTITLPCRICKII 322  
Db |||||  
QY 361 NRWQEVGKAMYAPPIRGQIRCSNITCLLLTRDGGKEISNTTEIFRPGGDMRDNRSEL 420  
Db |||||  
QY 323 NMWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGIN-ENGTEIFRPGGDMRDNRSEL 381  
Db |||||  
QY 421 YKVKVVKIEPLGVAPTKAKRRVVQREKRAVTLGAMFLGLGAAGSTMGARSLLTVQARQ 480  
Db |||||  
QY 382 YKVKVVKIEPLGVAPTKCKRRVVQREKRAVGIGAVFLGLGAAGSTMGASMTLTVOARL 441  
Db |||||  
QY 481 LLSGIVQQNNLLRAIEAQOHLQLTWGIK-LQARVLAVERYLKDQQLLGWGCSGKLI 539  
Db |||||  
QY 442 LLSGIVQQNNLLRAIEAQOQMLQLTWGIKLQARVLAVERYLGDQQLLGWGCSGKLI 501  
Db |||||  
QY 540 CTTAVPWNASWSNKSDDIWNNTWWEWEREIDNYTNLYTLIEESQNOEKNEQELLEL 599  
Db |||||  
QY 502 CTTAVPWNASWSNKSLLRIWNNTWWEWEREIDNYTSEIYTLIEESQNOEKNEQELLEL 561  
Db |||||  
QY 600 DKWASLWNNWFDISKWLWY 617  
Db |||||  
QY 562 DKWASLWNNWFDITNLWLY 579  
Db |||||

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:12:55 ; Search time 17.6158 Seconds  
(without alignments)  
3380.063 Million cell updates/sec

Title: US-09-891-609A-4  
Perfect score: 3327  
Sequence: 1 SAVEKLVTVYGVVWKEA.....DKWASLWNWFDISKWLWYIK 619  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2906	87.3	843	1 H44001	env polyprotein pr
2	2906	87.3	847	2 T09448	envelope glycoprot
3	2900.5	87.2	852	2 T12016	envelope glycoprot
4	2889	86.8	847	2 S13289	env protein - huma
5	2856	85.8	855	1 VCLJAJ2	env polyprotein pr
6	2851	85.7	856	1 VCLJH3	env polyprotein pr
7	2848.5	85.6	861	1 VCLJIV	env polyprotein pr
8	2839.5	85.3	851	2 S13985	env polyprotein pr
9	2838	85.3	856	1 VCLJVL	env polyprotein pr
10	2836	85.2	861	1 VCLJSC	env polyprotein pr
11	2830	85.1	854	2 S13288	env protein - huma
12	2817	84.7	852	1 VCLJBR	env polyprotein -
13	2813.5	84.6	868	1 VCLJH4	env polyprotein pr
14	2792.5	83.9	729	1 VCLJXK	env polyprotein pr
15	2792.5	83.9	861	1 VCLJKB	env polyprotein pr
16	2787.5	83.8	856	1 VCLJ3W	env polyprotein pr
17	2772.5	83.3	859	1 VCLJMN	env polyprotein pr
18	2593	77.7	855	1 VCLJZR	env polyprotein pr
19	2583.5	77.7	859	2 T01672	envelope polyprote
20	2571	77.3	853	2 S54384	envelope polyprote
21	2540.5	76.4	856	1 A44963	env polyprotein pr
22	2534.5	76.2	846	1 VCLJND	env polyprotein pr
23	2191	65.9	854	1 VCLJJS1	env polyprotein pr
24	1969.5	59.2	506	2 A40218	envelop glycoprote
25	1945.5	58.5	445	2 A41621	env polyprotein M
26	1941.5	58.4	443	2 C41621	env polyprotein P
27	1849	55.6	454	2 B41621	env polyprotein D
28	1839	55.3	495	2 S31493	env polyprotein -
29	1803	54.2	877	2 S49197	envelope protein p

ALIGNMENTS

RESULT 1

H44001  
env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)  
N;Alternate names: coat polyprotein  
N;Contains: coat protein gp120; coat protein gp41  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Sep-1994  
C;Accession: H44001  
R;Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
J. Virol. 66, 6587-6600, 1992  
A;Title: Complete nucleotide sequence, genome organization, and biological properties of  
A;Reference number: A44001; MUID:93021387; PMID:1404605  
A;Accession: H44001  
A;Molecule type: DNA  
A;Residues: 1-843 <liY>  
A;Cross-references: GB:M93258  
C;Genetics:  
A;Gene: env  
A;Superfamily: type E retrovirus env polyprotein  
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;19-35/Region: hydrophobic  
F;30-489/Product: coat protein gp120 #status predicted <GP1>  
F;490-843/Product: coat protein gp41 #status predicted <GP2>  
F;499-515/Region: hydrophobic  
F;673-689/Region: hydrophobic  
F;738-755/Domain: transmembrane #status predicted <TMN>  
F;87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435,

Query Match	87.3%	Score	2906;	DB 1;	Length	843;			
Best Local Similarity	84.2%	Pred. No.	1.4e-202;						
Matches	544;	Conservative	37;	Mismatches	35;	Indels	30;	Gaps	5;
QY	1	SAVEKLVTVYGVVWKEATTLFCASDAKAYDTEVHNVWATHACVTPDNPQETVLEN	60						
DB	28	SAAEQLVTVYGVVWKEATTLFCASDAKAYDTEVHNVWATHACVTPDNPQEVKLEN	87						
QY	61	VTFENFNNKNNMVQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATNTKSSNWKEM	120						
DB	88	VTFENFNNKNNMVQMHEDIISLWDQSLKPCVKLTPLCVTLNCTLRNATNTTSSSWETM	147						
QY	121	DRGEIKNCPSFKV-----GAGKLNCNTSVITQACPKVSFE	155						
DB	148	EXGEIKNCPSFNITTSIRDKVQKEYALFYNLDVVPIDNASYRLISCNTSVITQACPKVSFE	207						
QY	156	PIPIHYCAPAGFAILKCNCKFKNGSGPCTNVSTVQCTHGIRPVWSTQLLINGSABEGVV	215						
DB	208	PIPIHYCAPAGFAILKCNCKFKNGTGTCTNVSTVQCTHGIRPVWSTQLLINGSABEEIV	267						
QY	216	IRSENFDTNAKTIIVQLKESVEINCTRPNNNTRKSTIGPGAFAYATGDIIGDIOAHCN	275						

```
Db 268 IRSENFNNAKTIIVQLNESVWINCTRPNNNTRKSNIGPRALYTTGIIIGDIRQAHCN 327
Qy 276 ISEKKNNTLKQIVTKLQAFQ-NKTIIVFKQSGGDPPEIVMHSFNCGGFFFCVNSTOLF 334
Db 328 LSKTQWNTLEQIAIKLKEQFGNKKTIIFNPSSGGDPPEIVTHSFNCGGFFFCVNSTOLF- 386
Qy 335 STWNNTIGPNTNGTITLPCRIOIINRWQEVGKAMYAPPPIRQIIRCSNITGLLLTRDG 394
Db 387 -TWNDRKLNNTGNITLPCRIOIINRWQEVGKAMYAPPPIRQIIRCSNITGLLLTRDG 445
Qy 395 GKEISNTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARVVQREKRAVTLGA 454
Db 446 GKD-TNGTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARVVQREKRAVGLGA 504
Qy 455 MFLGFLGAAGSTMGARSITLTVQAROLLSGIVQOQNNLLRAIEAQQHLLQLTVMGK-LQ 513
Db 505 LFLGFLGAAGSTMGARSITLTVQAROLLSGIVQOQNNLLRAIEAQQHLLQLTVMGKIQ 564
Qy 514 ARVLAVERYLKDOQLLGWCSGKLICTTAVPWNASWSNKSLDQIWNNTMTMWEWEREIDN 573
Db 565 ARVLAVERYLRDQQLLGWCSGKLICTTVPNTSWNSKSLNEIWNNTMTMWEWEREIDN 624
Qy 574 YTNLYTLIESONQOEKNEQELLELDKWASLWNWFDISKWLWYIK 619
Db 625 YTHIIVSLIESONQOEKNEQELLELDKWASLWNWFDITKWLWYIK 670

RESULT 2
T09448
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: T09448
R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,
submitted to the EMBL Data Library, July 1996
A:Reference number: Z16673
A:Accession: T09448
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-847 <PAN>
A:Cross-references: EMBL:U63632; NID:g1465777; PID:g1465781
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match 87.3%; Score 2906; DB 2; Length 847;
Best Local Similarity 85.4%; Pred. No. 1.5e-202;
Matches 554; Conservative 25; Mismatches 38; Indels 32; Gaps 6;

Qy 1 SAVEKLWTVVYGVVPVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN 60
Db 28 SAVEKLWTVVYGVVPVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN 87
Qy 61 VTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTN-LKNATNTKSNWKE 120
Db 88 VTEHFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCKDV-NAINTTNDSEGTM 146
Qy 121 DRGEIKNCSFKVAG-----KLINCNTSVITQACPVS 153
Db 147 ERGEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPVS 206
Qy 154 FEPIPIHYCAPAGAILKNDKFNKSGPCTNVTSTVQCTHGIRPVVSTOLLNGLSABEG 213
Db 207 FEPIPIHYCAPAGAILKNDKFNKSGPCKNVTSTVQCTHGIRPVVSTOLLNGLSABEE 266
Qy 214 VVIRSENFNTDAKTIIVQLKESVEINCTRPNNNTRKSIITIGPRAFYATGDIIGDIRQAH 273
Db 267 VVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGIEIGDIRQAH 326
Qy 274 CNISGKKNNTLKQIVTKLQAFQGNKTIIVFKQSGGDPPEIVMHSFNCGGFFFCVNSTOLF 333
Db 327 CNISRAKNWDTLKQIVTKLREQFNKTIIVFNHSSGGDPPEIVMHSFNCGGFFFCVNSTOLF 386
```

```
Qy 334 NSTW-NNTIGPNNYNG-TITLPCRIOIINRWQEVGKAMYAPPPIRQIIRCSNITGLLLT 391
Db 387 NSTWNNTEGSSNTEGNTITLPCRIOIINRWQEVGKAMYAPPPIRQIIRCSNITGLLLT 446
Qy 392 RGGGKEISNTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARVVQREKRAVT 451
Db 447 RGGGIN-ENGTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARVVQREKRAVG 505
Qy 452 LGAMFLGFLGAAGSTMGARSITLTVQAROLLSGIVQOQNNLLRAIEAQQHLLQLTVMGK 511
Db 506 LGAVFLGFLGAAGSTMGARSITLTVQAROLLSGIVQOQNNLLRAIEAQQHLLQLTVMGK 565
Qy 512 -LQARVLAVERYLKDOQLLGWCSGKLICTTAVPWNASWSNKSLDQIWNNTMTMWEWERE 570
Db 566 LQARVLAVERYLRDQQLLGWCSGKLICTTAVPWNASWSNKSLDRIWNNTMTMWEWERE 625
Qy 571 IDNYTNLYTLIESONQOEKNEQELLELDKWASLWNWFDISKWLWYIK 619
Db 626 IDNYTSEIYTLIESONQOEKNEQELLELDKWASLWNWFDITKWLWYIK 674

RESULT 3
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12016
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A:Reference number: Z17379; MUID:98178716; PMID:9519894
A:Accession: T12016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <NCC>
A:Cross-references: EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g2351784
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match 87.2%; Score 2900.5; DB 2; Length 852;
Best Local Similarity 84.3%; Pred. No. 3.7e-202;
Matches 552; Conservative 32; Mismatches 32; Indels 39; Gaps 7;

Qy 1 SAVEKLWTVVYGVVPVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN 60
Db 28 SAAEQRLVTVVYGVVPVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN 87
Qy 61 VTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTN-LKNATNTKSNWKE 119
Db 88 VTENFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDYLNDTNTTSNNGG 147
Qy 120 MRGEIKNCSFKV-----GAGKLINCNTSVITQACPVS 152
Db 148 MEGGEIKNCSFNITTRIGNKVQKEYALFYKLDVVPIDNTTTSYRLINCNTSVITQACPVS 207
Qy 153 SEPEPIHYCAPAGAILKNDKFNKSGPCTNVTSTVQCTHGIRPVVSTOLLNGLSABEE 212
Db 208 SEPEPIHYCTTAPGAFALLKCKDKFNKSGPCTNVTSTVQCTHGIRPVVSTOLLNGLSABEE 267
Qy 213 GVVISSENFNTDAKTIIVQLKESVEINCTRPNNNTRKSIITIGPRAFYATGDIIGDIRQA 272
Db 268 EVVIRSENFNTDAKTIIVQLNESVEINCTRPNNNTRKSIHIGPQALYATGAIIGDIRQA 327
Qy 273 HCNISGKKNNTLKQIVTKLQAFQGNKTIIVFKQSGGDPPEIVMHSFNCGGFFFCVNSTQL 332
Db 328 HCNISRAKNWNTLKQIVKKLKVQFGNKTIFPNQSGGDPPEIVMHSFNCGGFFFCVNTKL 387
Qy 333 FNSTW--NNTIGPNNY-----NGTITLPCRIOIINRWQEVGKAMYAPPPIRQIIRCSN 385
Db 388 FNSTWNNFTNW--NDTEGTEGNTITLPCRIOIINRWQEVGKAMYAPPPIRQIIRCSN 445
Qy 386 TGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARVVOR 445
```

```
Db 446 TGLLLTRDGGTN-NSTNETFRPGGDMRDNWSELYKVKVIEPLGVAPTAKRRVQR 504
QY 446 EKRAVTLGAMFLGFLGAAGSTMGARSLTLTVQARQLLSGIVQOQNLLRAIEAQHLLQL 505
Db 505 EKRAVIGLALFLGFLGAAGSTMGAAAVTLTVQARQLLSGIVQOQNLLRAIEAQHLLQL 564
QY 506 TWGIIK-LQARVLAVERYLKQQLLIGWCSGKLICTTAVPNWNASWNSKSLDQIWNMTW 564
Db 565 TWGIIKQQLARVLAVERYLRDQQLLIGWCSGKLICTTAVPNWNASWNSKSLDQIWNMTW 624
QY 565 MEWEREIDNYTLIYTLIESQOQKEQELLELDKWASLWNPFDISKWLWYIK 619
Db 625 MEWEREIDNYTSLIYSLIESQOQKQELLELDKWASLWNPFDITKWLWYIK 679

RESULT 4
S13289
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S13289
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13289
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <OR>
C:Superfamily: type E retrovirus env polyprotein

Query Match 86.8%; Score 2889; DB 2; Length 847;
Best Local Similarity 84.9%; Pred. No. 2.5e-201;
Matches 551; Conservative 24; Mismatches 42; Indels 32; Gaps 6;

QY 1 SAVEKLWTVTVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPPNPQEVLEN 60
Db 28 SATEKLWTVTVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPPNPQEVLEN 87
QY 61 VTENFNWKNMVQEHEDIISLDQSLKPCVKLTPLCVTLCTNLKNAITNTKSSNWKE 120
Db 88 VTENFNWKNMVQEHEDIISLDQSLKPCVKLTPLCVTLCTNLKNAITNTKSSNWKE 146
QY 121 DRGEIKNCSFKVAG-----KLNCTNSVITQACPVS 153
Db 147 ERGEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNTSYRLISCDTSVITQACPVS 206
QY 154 FEPPIHYCAPAGFAILKNDKFNKSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEG 213
Db 207 FEPPIHYCAPAGFAILKNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEB 266
QY 214 VVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIIGGPRAFYATGDIIDIROAH 273
Db 267 VVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIIGGPRAFYATGDIIDIROAH 326
QY 274 CNISGKWNNTLKOIVTKLQAPGKNTIVFKQSSGGDPEIVMHSFNCGGEFFYCNSTQLP 333
Db 327 CNISRAKNDTLKOIVTKLREQENKTIIVFNHSSGGDPEIVMHSFNCGGEFFYCNSTQLN 386
QY 334 NSTW-NNTIGPNNTNG-TITLPCRIKOIINRWQEVKAMVAPPRIQIRCSSNITGLLT 391
Db 387 NSTWNNTEGSNNTEGNTITLPCRIKOIINRWQEVKAMVAPPRIQIRCSSNITGLLT 446
QY 392 ROGKEISNTTEIFRPGGDMRDNWSELYKVKVIEPLGVAPTAKRRVQREKRAVT 451
Db 447 RDGGIN-ENGTEIFRPGGDMRDNWSELYKVKVIEPLGVAPTAKRRVQREKRAVG 505
QY 452 LGAMFLGFLGAAGSTMGARSLTLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIIK 511
Db 506 IGAFLGFLGAAGSTMGARSLTLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIIK 565
QY 512 -LQARVLAVERYLKQQLLIGWCSGKLICTTAVPNWNASWNSKSLDQIWNMTWMEWERE 570
```

```
Db 566 QLQARVLAVERYLQQLLIGWCSGKLICTTAVPNWNASWNSKSLDQIWNMTWMEWERE 625
QY 571 IDNYTNTLVTLESQOQKEQELLELDKWASLWNPFDISKWLWYIK 619
Db 626 IDNYTSEIYTLIESQOQKEQELLELDKWASLWNPFDITKWLWYIK 674

RESULT 5
VCLJ2A2
env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: Host Homo sapiens (man)
C>Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03976
R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-Shi
Science 227, 484-492, 1985
A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).
A:Reference number: A04003; MUID:85090453; PMID:2578227
A:Accession: A03976
A:Molecule type: DNA
A:Residues: 1-855 <SN>
A:Cross-references: GB:K02007; NID:G328658; PIDN:AAB59882.1; PID:G328666
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-509/Product: exterior membrane glycoprotein #status predicted <EXT>
F:510-855/Product: transmembrane glycoprotein #status predicted <TM>
F:87,129,140,150,184,190,200,244,265,292,298,304,334,341,350,364,388,394,400,408,445,458,
F:610,624,636,815/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 85.8%; Score 2856; DB 1; Length 855;
Best Local Similarity 82.3%; Pred. No. 6.2e-199;
Matches 540; Conservative 32; Mismatches 46; Indels 38; Gaps 6;

QY 1 SAVEKLWTVTVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPPNPQEVLEN 60
Db 28 SATEKLWTVTVYGVVPWKAEATTLFCASDAKAYDTEVHNWATHACVPTDPPNPQEVLEN 87
QY 61 VTENFNWKNMVQEHEDIISLDQSLKPCVKLTPLCVTLCTNLKNAITNTKSSNWKE 120
Db 88 VTENFNWKNMVQEHEDIISLDQSLKPCVKLTPLCVTLCTNLKNAITNTKSSNWKE 147
QY 121 DRGEIKNCSFKVAG-----KLNCTNSVITQACPVS 148
Db 148 IRGEIKNCSFNITTSIRDKIQKENALFRNLDVVPIDNNSTNTYRLIHCNRSVITQACPVS 207
QY 149 CPKVSFEPPIHYCAPAGFAILKNDKFNKSGPCTNVSTVQCTHGIRPVVSTQLLNGS 208
Db 208 CPKVSFEPPIHYCAPAGFAILKNDKFNKSGPCTNVSTVQCTHGIRPVVSTQLLNGS 267
QY 209 LAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIIGGPRAFYATGDIID 268
Db 268 LAEEGVVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIIGGPRAFYATGDIID 327
QY 269 IRQAHCHNISGKWNNTLKOIVTKLQAPGKNTIVFKQSSGGDPEIVMHSFNCGGEFFYC 327
Db 328 IRQAHCHNISGKWNNTLKOIVTKLQAPGKNTIVFKQSSGGDPEIVMHSFNCGGEFFYC 387
QY 328 NSTQLFNSTW-NNTIGPNNTNGTITLPCRIKOIINRWQEVKAMVAPPRIQIRCSSNI 385
Db 388 NSTQLFNSTW-NNTIGPNNTNGTITLPCRIKOIINRWQEVKAMVAPPRIQIRCSSNI 446
QY 386 TGLLLTRDGGKEISNTTEIFRPGGDMRDNWSELYKVKVIEPLGVAPTAKRRVQREKRAVT 445
Db 447 TGLLLTRDGGTNTVNDTEVFRPGGDMRDNWSELYKVKVIEPLGVAPTAKRRVQREKRAVG 506
QY 446 EKRAV-TLGAMFLGFLGAAGSTMGARSLTLTVQARQLLSGIVQOQNLLRAIEAQHLLQL 504
Db 507 EKRAVIGVAMFLGFLGAAGSTMGARSLTLTVQARQLLSGIVQOQNLLRAIEAQHLLQL 566
```

```
QY 505 LTVWGIK-LQARVLAVERYLKDOQLLGIWCGSKLCTTAVPWNASWSNKSLSDDIWNWNT 563
Db 567 LTVWGIK-LQARVLAVERYLKDOQLLGIWCGSKLCTTAVPWNASWSNKSLSDDIWNWNT 626
QY 564 WMEWEREIDNTNLIYTLIEESQOQKNEQELLELDKWSLWNNWFDISKWLWYIK 619
Db 627 WMEWEREIDNTNLIYTLIEESQOQKNEQELLELDKWSLWNNWFDISKWLWYIK 682

RESULT 6
VCLJH3
env polypeptide precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)
N:Alternate names: coat polypeptide
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03973
R:Ratner, L.; Haeltline, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dora-
nberger, J.A.; Papas, T.S.; Grayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, P.
Nature 313, 277-284, 1985
A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
A:Reference number: A93353; MUID:85111123; PMID:2578615
A:Accession: A03973
A:Molecule type: DNA
A:Residues: 1-856 <BAT>
A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:g326383; PIDN:AAA442
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypep-
F:1-30/Domain: signal sequence #status predicted <SIG>
F:1-511/Product: exterior membrane glycoprotein #status predicted <EXT>
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,303,332,339,356,386,392,397,406
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 85.7%; Score 2851; DB 1; Length 856;
Best Local Similarity 83.4%; Pred. No. 1.4e-198;
Matches 548; Conservative 32; Mismatches 37; Indels 40; Gaps 10;

QY 1 SAVEKLWTVYVGVVWKEATTLFCASDAKAYDTEHNVWATHACVPTDNPQEVVLN 60
Db 29 SATEKLWTVYVGVVWKEATTLFCASDAKAYDTEHNVWATHACVPTDNPQEVVLN 88
QY 61 VTENFNWKNMVEQMHEDIISLWDSQKPCVKLTPLCVLHCTNLKNAATNTKSSNKE- 119
Db 89 VTENFNWKNMVEQMHEDIISLWDSQKPCVKLTPLCVLHCTNLKNAATNTKSSNKE- 148
QY 120 MDRGEIKNCSFKVGA---GK-----LINCNTSVITQACPKV 152
Db 149 MEKGEIKNCSFNISTSRGKVQKEYAFYKLDIIPDNDTTSVTLTSCNTSVITQACPKV 208
QY 153 SFPEPIHYCAPAGFAILKCNDDKFNKSGPCTNVSTVQCTHGRPVVSTQLLNGSLAE 212
Db 209 SFPEPIHYCAPAGFAILKCNNTKFNKSGPCTNVSTVQCTHGRPVVSTQLLNGSLAE 268
QY 213 GVVISNFTDNAKTIIVQLKESVEINCTRPNNTRKSIIT--GPGRAFVATGDI 270
Db 269 EVVISNFTDNAKTIIVQLNQSVEINCTRPNNTRKSIIRIQRGPRAVFTIGK-IGNMR 327
QY 271 QAHCNTSGEKNWNLKQIVTLKQAFG-NKTIIVKQSSGDPPELVHSPNCGGEPFCNS 329
Db 328 QAHCNTSRKAWNNTLKQIDSKLREQFGNKKTIIFKQSSGDPPELVHSPNCGGEPFCNS 387
QY 330 TQLFNSTWNN---TIGPNTNG--TITLPCRKQIINRWQEVGKAMYAPPIRGICSS 383
Db 388 TQLFNSTWNTWSTKSNNTESDITLPCRKQIINRWQEVGKAMYAPPIRGICSS 447
QY 384 NITGLLLTRDGGKEISNTTIFRPGGDMRDNWSELYKYKVKVIEPLGVAPTAKRRVV 443
Db 448 NITGLLLTRDGGNS--NNESEIFRPGGDMRDNWSELYKYKVKVIEPLGVAPTAKRRVV 506
```

```
QY 444 QREKAVLTGMFLGAGSTMGARSLTLTVQARQLLSGIVQOQNNLLRAIEAQHLL 503
Db 507 QREKAVGTGALFLGAGSTMGAASWTLTVQARQLLSGIVQOQNNLLRAIEAQHLL 566
QY 504 QUTVWGIK-LQARVLAVERYLKDOQLLGIWCGSKLCTTAVPWNASWSNKSLSDDIWNWNT 562
Db 567 QUTVWGIK-LQARVLAVERYLKDOQLLGIWCGSKLCTTAVPWNASWSNKSLSDDIWNWNT 626
QY 563 WMEWEREIDNTNLIYTLIEESQOQKNEQELLELDKWSLWNNWFDISKWLWYIK 619
Db 627 WMEWEREIDNTNLIYTLIEESQOQKNEQELLELDKWSLWNNWFDISKWLWYIK 683

RESULT 7
VCLJLV
env polypeptide precursor - human immunodeficiency virus type 1 (isolate LAV-la)
N:Alternate names: coat polypeptide
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03975
R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.
Cell 40, 9-17, 1985
A:Title: Nucleotide sequence of the AIDS virus, LAV.
A:Reference number: A90866; MUID:85099333; PMID:2981635
A:Accession: A03975
A:Molecule type: DNA
A:Residues: 1-861 <WAI>
A:Cross-references: GB:K02013; NID:g326417; PIDN:AAB59751.1; PID:g326424
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypep-
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>
F:517-861/Product: transmembrane glycoprotein #status predicted <TM>
F:88,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411,
F:616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predic
F:616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 85.6%; Score 2848.5; DB 1; Length 861;
Best Local Similarity 82.8%; Pred. No. 2.2e-198;
Matches 548; Conservative 30; Mismatches 39; Indels 45; Gaps 10;

QY 1 SAVEKLWTVYVGVVWKEATTLFCASDAKAYDTEHNVWATHACVPTDNPQEVVLN 60
Db 29 SATEKLWTVYVGVVWKEATTLFCASDAKAYDTEHNVWATHACVPTDNPQEVVLN 88
QY 61 VTENFNWKNMVEQMHEDIISLWDSQKPCVKLTPLCVLHCTNLKNAATNTKSSNKE- 119
Db 89 VTENFNWKNMVEQMHEDIISLWDSQKPCVKLTPLCVLHCTNLKNAATNTKSSNKE- 148
QY 120 MDRGEIKNCSFKVGA---GK-----LINCNTSVITQ 147
Db 149 MEKGEIKNCSFNISTSRGKVQKEYAFYKLDIIPDNDTTSVTLTSCNTSVITQ 208
QY 148 APCKVSFEPIHYCAPAGFAILKCNDDKFNKSGPCTNVSTVQCTHGRPVVSTQLLNG 207
Db 209 APCKVSFEPIHYCAPAGFAILKCNNTKFNKSGPCTNVSTVQCTHGRPVVSTQLLNG 268
QY 208 SLAEGVIRSENFTDNAKTIIVQLKESVEINCTRPNNTRKSIIT--GPGRAFVATGDI 265
Db 269 SLAEEVWIRSANFTDNAKTIIVQLNQSVEINCTRPNNTRKSIIRIQRGPRAVFTIGK- 327
QY 266 IGDROAHCHNISEKKNWNLKQIVTLKQAFG-NKTIIVKQSSGDPPELVHSPNCGGEP 324
Db 328 IGMROAHCHNISRKAWNNTLKQIDSKLREQFGNKKTIIFKQSSGDPPELVHSPNCGGEP 387
QY 325 FYCNSQLFNSTWNN---TIGPNTNG--TITLPCRKQIINRWQEVGKAMYAPPIRG 378
Db 388 FYCNSQLFNSTWNTWSTKSNNTESDITLPCRKQIINRWQEVGKAMYAPPIRG 447
QY 379 IRCSSNITGLLLTRDGGKEISNTTIFRPGGDMRDNWSELYKYKVKVIEPLGVAPTAK 438
```





```
Db 567 TVWGIKQARILAVERYLKDOQLGIGWCGSKLICCTTAVPMNASNSKLEQIWNMTW 626
Qy 565 MEWEIEDNTNLIYTLIESNQKNEQELLELDKWASLWVFNFDISKWLWYIK 619
Db 627 MEWDREINNTYSLIHSLESNQKNEQELLELDKWASLWVFNITNWLWYIK 681

RESULT 12
VCUJBR
env polyprotein - human immunodeficiency virus type 1 (isolate BR)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Feb-1997
C;Accession: A31667
R;Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S
Virology 168, 79-89, 1989
A;Title: Biological and molecular characterization of human immunodeficiency virus (HIV-1)
A;Reference number: A94389; MUID:89085613; PMID:2789516
A;Accession: A31667
A;Molecule type: DNA
A;Residues: 1-852 <ANA>
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-516/Product: coat protein gp120 #status predicted <CP1>
F;517-852/Product: coat protein gp41 #status predicted <CP2>

Query Match 84.7%; Score 2817; DB 1; Length 852;
Best Local Similarity 81.6%; Pred. No. 4.1e-196;
Matches 533; Conservative 38; Mismatches 46; Indels 36; Gaps 7;

Qy 1 SAVEKLVVTVYGVVWKKEATTLFCASDAKAYDTEHNVWATHACVPTDNPQEVLEN 60
Db 29 SATDKLVVTVYGVVWKKEATTLFCASDAKAYDTEHNVWATHACVPTDNPQEVLEN 88
Qy 61 VTENFNNKNNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLTNKNATNTSSNWKEM 120
Db 89 VTENFNNKNNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLTNKNATNTSSNWKEM 147
Qy 121 DRGEIKNCSFKVAG-----KLNKNTSVITQACP 151
Db 148 EGGEKNCNCSFNITTSIRDKMKQKEYALFYKLDIVPIDNDKNTNRYLISCNSTVITQACP 207
Qy 152 VSPPEPIHYCAPAGFAILLKCNKKFNGSGPCNTVSTVQCTHGIRPVVSTQALLNGSLAE 211
Db 208 VTEPIPIHYCAPAGFAILLKCNKKFNGSGPCNTVSTVQCTHGIRPVVSTQALLNGSLAE 267
Qy 212 EGVVIRSEFTDNAKTIIVQLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIR 271
Db 268 EGVVIRSEFTNNKTIIVQLNESVEINCTRPNNTRKRIITMGPRVYVYTTGQIIGDIR 327
Qy 272 AHCNIISGEKNNTLKQIVTKLQAOFGNKTIVFKQSSGGDPEIVMHSFNCGGEFFYNSTQ 331
Db 328 AHCNLSRSKWENTLKQIVTKLRVQFNKTIIVFNRSGGDPEIVMHSFNCGGEFFCNTQ 387
Qy 332 LFNSTW-NTTIGPNNNGT--ITLPCRKIQINRWQEVKAMVAPTRIGQIRCSSNITGL 388
Db 388 LFNSTWYRNTTG-NITEGNSPITLPCRKIQIINRWQEVKAMVAPTRIGQIRCSSNITGL 446
Qy 389 LLTRDGGKEISNT-TEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTKAKERVVQREK 447
Db 447 LLTRDGGNNNETTDEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTKAKERVVQREK 506
Qy 448 RAVTLGAMFLGFGAAGSTWGARSLLTIVQAROLLSGIVQOQNLLRAIEAQHLLQLTV 507
Db 507 RAVGLGALFLGFLGAAGSTWGAAASLLTIVQARLLSGIVQOQNLLMAIEAQHMLLTV 566
Qy 508 WGIK-LOARVLAVERYLKDOQLGIGWCGSKLICCTTAVPMNASNSKSLDQIWNMTWME 566
Db 567 WGIKQIARVLAVERYLKDOQLGIGWCGSKLICCTTAVPMNASNSKSLDIWNMTWME 626
```

```
Qy 567 WEIEDNTNLIYTLIESNQKNEQELLELDKWASLWVFNFDISKWLWYIK 619
Db 627 WEIEDNTNLIYSLIEDSQIQEKNEKELLELDKWASLWVFNITNWLWYIK 679

RESULT 13
VCUJH4
env polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C;Accession: C25523
R;Desai, S.M.; Kalyanaram, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devare, S
Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986
A;Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human
A;Reference number: A94136; MUID:87041461; PMID:3490666
A;Accession: C25523
A;Molecule type: DNA
A;Residues: 1-868 <DES>
A;Cross-references: GB:M13137; NID:g326460; PIDN:AAA44311.1; PID:g326467
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-521/Product: coat protein gp120 #status predicted <GP1>
F;522-868/Product: coat protein gp41 #status predicted <GP2>
F;89,138,139,142,166,195,198,208,245,252,273,287,300,312,342,349,365,371,395,405,409,459

Query Match 84.6%; Score 2813.5; DB 1; Length 868;
Best Local Similarity 79.3%; Pred. No. 7.6e-196;
Matches 533; Conservative 36; Mismatches 44; Indels 59; Gaps 8;

Qy 1 SAVEKLVVTVYGVVWKKEATTLFCASDAKAYDTEHNVWATHACVPTDNPQEVLEN 60
Db 30 SAANKLVVTVYGVVWKKEATTLFCASDAKAYDTEHNVWATHACVPTNPNQEVLEN 89
Qy 61 VTENFNNKNNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCT--NLKNATNTKSSN-- 116
Db 90 VTENFNNKNNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNTNTNTTTELSII 149
Qy 117 --WKMDRGEIKNCSFKVAG-----KLNKNT 142
Db 150 VVMEQRKGKGMNCSFNITTSIRDKVQREYALFYKLDVPEIDDNKNTNTNTKRYLINCNT 209
Qy 143 SVITQACPVSPEPIPIHYCAPAGFAILLKCNKKFNGSGPCNTVSTVQCTHGIRPVVSTQ 202
Db 210 SVITQACPVSPEPIPIHYCTPTGFAILLKCNKKFNGSGPCNTVSTVQCTHGIRPVVSTQ 269
Qy 203 LLNGLSLABEGVVIRSEFTDNAKTIIVQLKESVEINCTRPNNTRKSIITIGPGRAPYAT 262
Db 270 LLNGLSLABEEVVIRSEFTNNAKTIIVQLNVSVENCTRPNNTRKRVTLGPRVWYTT 329
Qy 263 GDIIGDIRQAHNCNISEKNNTLKQIVTKLQAOFGNKTIVFKQSSGGDPEIVMHSFNCGG 322
Db 330 GEILGNIRQAHNCNISRAQWNNTLQIATTLREQFGNKTIAFNQSSGGDPEIVMHSFNCGG 389
Qy 323 EEPYCNSTOLFNSWNTTIGPNNNGT-----ITLPCRKIQIINRWQEVGKA 369
Db 390 EEPYCNSTOLFNSAWNT-----SNGTWSVTRKQKDTGDIITLPCRKIQIINRWQVVGKA 444
Qy 370 MYAPPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVKVIE 429
Db 445 MYALPIKGLIRCSSNITGLLLTRDGGGE-NQTEIFRPGGDMRDNRSELYKYKVKVIE 503
Qy 430 PLGVAPTKAKRRVVQREKRAV-TLGAMFLGFGAAGSTWGARSLLTIVQAROLLSGIVQ 488
Db 504 PLGVAPTKAKRRVVQREKRAVGLGAMFLGFGAAGSTWGATSMALTIVQAROLLSGIVQ 563
Qy 489 QNNLLRAIEAQHLLQLTVWGIK-LOARVLAVERYLKDOQLGIGWCGSKLICCTTAVPMN 547
Db 564 QNNLLRAIEAQHLLQLTVWGIKQIARVLAVERYLKDOQLGIGWCGSKLICCTTAVPMN 623
```



```

Qy 548 ASNSKSLDQIQWNNMTMEWEREIDNTNLIYTLIESQOQEKQOELLJELDKWASLWN 607
Db 624 ASNSKTLDDQIQWNNMTMEWDREIDNTNLIYTLIESQOQEKQOELLJELDKWASLWT 683
Qy 608 WFDISKWLWYIK 619
Db 684 WSDITKWLWYIK 695

RESULT 14
VCLUXX
env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp32)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp32
C;Species: human immunodeficiency virus type 1, HIV-1
C;Notes: host Homo sapiens (man)
A;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Mar-1996
C;Accession: B42395
R;Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.
Virology 189, 534-546, 1992
A;Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated
A;Reference number: A42995; MUID:92351552; PMID:1322587
A;Accession: B42995
A;Molecule type: mRNA
A;Residues: 1-729 <SHI>
A;Cross-references: GB:S41266; GB:D01206
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-689/Domain: extracellular #status predicted <EXT>
F;1-33/Domain: signal sequence #status predicted <SIG>
F;17-33/Region: hydrophobic #status predicted
F;34-517/Product: coat protein gp120 #status predicted <CP1>
F;514-517/Region: cleavage processing #status predicted
F;518-729/Product: coat protein gp32 #status predicted <CP2>
F;518-534/Region: hydrophobic #status predicted
F;690-711/Domain: transmembrane #status predicted <TM1>
F;712-729/Domain: intracellular #status predicted <IN1>
F;93,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602,603,604,605,606,607,608,609,610,611,612,613,614,615,616,617,618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,637,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,656,657,658,659,660,661,662,663,664,665,666,667,668,669,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,738,739,740,741,742,743,744,745,746,747,748,749,750,751,752,753,754,755,756,757,758,759,760,761,762,763,764,765,766,767,768,769,770,771,772,773,774,775,776,777,778,779,780,781,782,783,784,785,786,787,788,789,790,791,792,793,794,795,796,797,798,799,800,801,802,803,804,805,806,807,808,809,810,811,812,813,814,815,816,817,818,819,820,821,822,823,824,825,826,827,828,829,830,831,832,833,834,835,836,837,838,839,840,841,842,843,844,845,846,847,848,849,850,851,852,853,854,855,856,857,858,859,860,861,862,863,864,865,866,867,868,869,870,871,872,873,874,875,876,877,878,879,880,881,882,883,884,885,886,887,888,889,890,891,892,893,894,895,896,897,898,899,900,901,902,903,904,905,906,907,908,909,910,911,912,913,914,915,916,917,918,919,920,921,922,923,924,925,926,927,928,929,930,931,932,933,934,935,936,937,938,939,940,941,942,943,944,945,946,947,948,949,950,951,952,953,954,955,956,957,958,959,960,961,962,963,964,965,966,967,968,969,970,971,972,973,974,975,976,977,978,979,980,981,982,983,984,985,986,987,988,989,990,991,992,993,994,995,996,997,998,999,1000

Query Match 83.9%; Score 2792.5; DB 1; Length 729;
Best Local Similarity 79.3%; Pred. No. 2e-194;
Matches 525; Conservative 41; Mismatches 47; Indels 49; Gaps 7;

Qy 1 SAVEKLWTVVYGVVPMKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
Db 34 SAAEQLWTVVYGVVPMKEATTTTLFCASDAKAYDTEAHNVWATHACVPTDPNPQEVVLN 93
Qy 61 VTENFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKGNATNTKSNWKM 120
Db 94 VTENFNWKNWVQMHENIISLWDQSLKPCVKLTPLCVTLHCTDLNRNTNNSSIBEKM 153
Qy 121 DRGEEKNCSPKVGAG-----KLI NCNTSVITQAC 149
Db 154 -KGEEKNCSPVTTNIRDVKQVEYALFYKLDVDPIDNNSNTCYRLISCDTSVITQAC 212
Qy 150 PKVSFEPIPIHYCAPAGFAILKCKNDKFPNGSGPCTNVSTVQCTHGIRPVVSTQLLNGSL 209
Db 213 PKVSFEPIPIHYCTPAGFALLKCNKTFNGTGPCKNVSTVQCTHGIRPVVSTQLLNGSL 272
Qy 210 AEEGVVIRSENFNTAKTIIIVOLKESVEINCRNNTRKSIITGPGRAFYATGDITGDI 269
Db 273 AEEGVVIRSENFNTDNVKTIIIVQLNETVKINCRNNTRKRVTMGPRVYVTTGEIIGDI 332
Qy 270 ROAHCNTSGEKNWNTLKOIVTKLOAQFGNKTIIVFKQSGGGDPPEIVMHSFNCGGFFFCNS 329
Db 333 ROAHCNISRAEWNTKLEOIANKLRRQFNKTIIVFNQSGGGDPPEIVMHNFCNGGFFFCYDS 392
Qy 330 TQLENST-----WNNTIGPNTNGTITLPCRKIQIINRWQEVGKAMYAPPIRGQIRCS 382
Db 393 SOLFNSTHLSNGTWNGT-GEEN---ITLPCRKIQIIVNMWQEVGKAMYAPPIRGQIRCS 447

```





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:03:35 ; Search time 10.7652 Seconds  
(without alignments)  
2994.034 Million cell updates/sec

Title: US-09-891-609A-4  
Perfect score: 3327  
Sequence: 1 SAVEKLVTVYGVVWKEA.....DKWASLWNWFDISKWLWYIK 619

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3275	98.4	847	1 ENV_HV1S1	P19550 human immun
2	2906	87.3	843	1 ENV_HV1Y2	P35961 human immun
3	2880.5	86.3	856	1 ENV_HV1S3	P05878 human immun
4	2870.5	86.3	848	1 ENV_HV1JR	P20871 human immun
5	2856	85.8	855	1 ENV_HV1A2	P03378 human immun
6	2851	85.7	856	1 ENV_HV1A1	P03375 human immun
7	2848.5	85.6	861	1 ENV_HV1BR	P03377 human immun
8	2839	85.3	865	1 ENV_HV1RH	P04579 human immun
9	2834	85.2	856	1 ENV_HV1PV	P03376 human immun
10	2834	85.2	856	1 ENV_HV1H2	P04578 human immun
11	2829.5	85.0	851	1 ENV_HV1B8	P04580 human immun
12	2820	84.8	852	1 ENV_HV1BN	P12488 human immun
13	2818	84.7	855	1 ENV_HV1OY	P20888 human immun
14	2813.5	84.6	868	1 ENV_HV1C4	P05879 human immun
15	2808	84.4	856	1 ENV_HV1H3	P04624 human immun
16	2807.5	84.4	852	1 ENV_HV1S3	P19549 human immun
17	2807	84.4	853	1 ENV_HV1MF	P19551 human immun
18	2803	84.3	856	1 ENV_HV1LW	P06226 human immun
19	2802	84.2	867	1 ENV_HV1J3	P12489 human immun
20	2792.5	83.9	861	1 ENV_HV1XB	P31819 human immun
21	2789	83.8	847	1 ENV_HV1W2	P05880 human immun
22	2787.5	83.8	856	1 ENV_HV1W1	P31872 human immun
23	2786	83.7	856	1 ENV_HV1MN	P05877 human immun
24	2593	77.9	855	1 ENV_HV1Z6	P04580 human immun
25	2583.5	77.7	859	1 ENV_HV1MA	P05882 human immun
26	2578	77.5	863	1 ENV_HV1Z8	P05882 human immun
27	2571	77.3	853	1 ENV_HV1Z2	P12487 human immun
28	2540.5	76.4	856	1 ENV_HV1ZH	P05881 human immun
29	2534.5	76.2	846	1 ENV_HV1ND	P18799 human immun
30	2513	75.5	853	1 ENV_HV1EL	P04581 human immun
31	2191	65.9	854	1 ENV_SIVCZ	P17281 chimpanzee
32	1585.5	47.7	421	1 ENV_HV1N5	P12490 human immun
33	1546.5	46.5	460	1 ENV_HV1Z3	P12491 human immun

ALIGNMENTS

RESULT 1				
ENV_HV1S1	ENV_HV1S1	STANDARD;	PRT;	847 AA.
AC	P19550;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).			
OC	Viruses; Retroviridae; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11691;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90347835; PubMed=2384920;			
RA	Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;			
RT	"Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";			
RL	J. Virol. 64:4390-4398 (1990).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M65024; AAA45072.1; -			
DR	PDB; 1OBE; 15-MAY-97.			
DR	HIV; M38428; ENVSSF162.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure.			
FT	SIGNAL	1	29	
FT	CHAIN	30	502	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	503	847	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	53	73	BY SIMILARITY.
FT	DISULFID	118	203	BY SIMILARITY.
FT	DISULFID	125	194	BY SIMILARITY.
FT	DISULFID	130	155	BY SIMILARITY.
FT	DISULFID	216	245	BY SIMILARITY.
FT	DISULFID	226	237	BY SIMILARITY.
FT	DISULFID	294	328	BY SIMILARITY.
FT	DISULFID	374	435	BY SIMILARITY.
FT	DISULFID	381	408	BY SIMILARITY.
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	135	135	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	232	232	N-LINKED (GLCNAC. . .) (POTENTIAL).

Q02837 simian immu  
P18040 human immu  
P20872 human immu  
P27977 simian immu  
P27757 simian immu  
P32536 human immu  
P24105 human immu  
P05883 human immu  
P05886 simian immu  
P18094 human immu  
P17755 human immu  
P04577 human immu



FT CARBOHYD 598 598 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 803 803 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

Query Match 87.3%; Score 2906; DB 1; Length 843;  
 Best Local Similarity 84.2%; Pred. No. 1.5e-217;  
 Matches 544; Conservative 37; Mismatches 35; Indels 30; Gaps 5;

Qy 1 SAVEKLWTVYGVVWKEATTTLCASDAKAYDTEVHNVWATHACVPTDPNPOEIVLEN 60  
 Db 28 SAAEQWLVTVYGVVWKEATTTLCASDAKAYDTEVHNVWATHACVPTDPNPOEIVLEN 87  
 Qy 61 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAITKSSNWKEM 120  
 Db 88 VTENFNWKNMVEQMHEDIISLDOSLKPCKVLTPLCVTLHCTNLKNAITKSSNWKEM 147  
 Qy 121 DRGEIKNCSPKV-----GAGKLNCSNTSVITQACPKVSFE 155  
 Db 148 EKEIKNCSPNITTSIRDKVQKEYALFVNDVVPIDNASYRLISCSNTSVITQACPKVSFE 207  
 Qy 156 PIPIHYCAPAGFAILLKNDKKGSGPCTNVSTVQCTHGRPVVSTOLLNGLSABEGVV 215  
 Db 208 PIPIHYCAPAGFAILLKNDKKGSGPCTNVSTVQCTHGRPVVSTOLLNGLSABEGVV 267  
 Qy 216 IRSENFDTNAKTIIVOLKESVEINCTPNNNTRKSIITIGPGRFYATGDIIGDIOAHCN 275  
 Db 268 IRSENFDTNAKTIIVOLKESVEINCTPNNNTRKSIINIGPGRFYATGDIIGDIOAHCN 327  
 Qy 276 ISGEKNNTLKQIVTKLQAFQ-NKTIIVFKQSGGDPPEIVMHSFNCGGFFYCNSTOLF 334  
 Db 328 LSKTOWENTLEQIAIKLKEQFGNNKTIIFNPSSGGDPPEIVMHSFNCGGFFYCNSTOLF- 386  
 Qy 335 STWNTIGNNTGTITLPCRIKQIINRMOEVCAMYPAPIRGOIRCSNITGLLITRDG 394  
 Db 387 -TWNTIRKLNNTGRNITLPCRIKQIINRMOEVCAMYPAPIRGOIRCSNITGLLITRDG 445  
 Qy 395 GKESINTEIFRGGGDMRDNWSELYKYKVKVIEPLGVAPTAKRVVQREKRAVTLGA 454  
 Db 446 GKD-TNGTEIFRGGGDMRDNWSELYKYKVKVIEPLGVAPTAKRVVQREKRAVTLGA 504  
 Qy 455 MFGFGLGAAGSTWGAASLTITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGKIK-LQ 513  
 Db 505 LFLGFLGAAGSTWGAASLTITVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGKIKQLQ 564  
 Qy 514 ARVLAVERYLKDOQLLIGWCSGKLCTTAVPWNASHNSKSLDQINWNTWMEWEREIDN 573  
 Db 565 ARVLAVERYLRDQQLLIGWCSGKLCTTAVPWNASHNSKSLDQINWNTWMEWEREIDN 624  
 Qy 574 YTMILYTLIESQNOQEKNEQELLEDKWSLWNPFDISKWLWYIK 619  
 Db 625 YTHIYSLIESQNOQEKNEQELLEDKWSLWNPFDISKWLWYIK 670

RESULT 3  
 ENV HV1SC  
 ID ENV HV1SC STANDARD; PRT; 856 AA.  
 AC P05878;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide gp160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).  
 OC Viruses; Retrod viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88219542; PubMed=3369091;  
 RA Gurgu C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,

RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;  
 "Envelope sequences of two new United States HIV-1 isolates";  
 Virology 164:531-536(1988).  
 RL -I- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN  
 CC 1984 IN SOUTHERN CALIFORNIA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 CC EMBL; M17450; -; NOT\_ANNOTATED\_CDS.  
 DR HIV; M17450; ENV\$SC.  
 DR InterPro: IPR000328; Env GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 510 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 511 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT SITE 760 760 IN-FRAME TERMINATION CODON.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 206 BY SIMILARITY.  
 FT DISULFID 125 197 BY SIMILARITY.  
 FT DISULFID 130 160 BY SIMILARITY.  
 FT DISULFID 219 247 BY SIMILARITY.  
 FT DISULFID 228 239 BY SIMILARITY.  
 FT DISULFID 296 330 BY SIMILARITY.  
 FT DISULFID 376 439 BY SIMILARITY.  
 FT DISULFID 383 412 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97055 MW; DAF4DA600BBA7A08 CRC64;

Query Match 86.6%; Score 2880.5; DB 1; Length 856;  
 Best Local Similarity 83.3%; Pred. No. 1.5e-215;  
 Matches 547; Conservative 31; Mismatches 40; Indels 39; Gaps 7;

```
QY 1 SAVEKLVVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
Db 28 SAAEQKLVVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 87
QY 61 VTENFNWKNMVEQMHEDIISLWDSKPCVKLTPLCVTLHCTNLK-----NATNTKGS 115
Db 88 VTENFNWKNMVEQMHEDIISLWDSKPCVKLTPLCVTLHCTNLKRDSTNATNTSS 147
QY 116 NWKEMDRGEIKNCSFKVGAG-----KLINCVTSVITQACP 150
Db 148 NRGKMEGTEGNSFNITTSIRSKVQKEYALFYKLDVVPIDNTSYTLINCVTSVITQACP 207
QY 151 KVSFEPIPIHYCAPAGFAILKNDKFPNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLA 210
Db 208 KVSFEPIPIHYCA-RWFAILNCKNKFNGPCTNVSTVQCTHGIRPVVSTHLLNGSLA 266
QY 211 EEGVIRSENFTDNAKTIIVQLKESVEINCTRNPNNTTRSIHIGPGRFATGDIIGDIR 270
Db 267 EEEVLRSENFTDNAKTIIVQLKEAVEINCTRNPNNTTRSIHIGPGRFATGDIIGDIR 326
QY 271 QAHCNISGEKNNTLKOIVTKLQAFQNGKTIIVFKQSGGDPPEIVMHSFNCGGFFVCNST 330
Db 327 QAHCNISRAKNNTLKOIVTKLQAFQNGKTIIVFNRSKGGDPPEIVMHSFNCGGFFVCNST 386
QY 331 QLFNSTWNTTIGPNTNG--TITLPCRIKOINRWQEVGKAMVAPPIRGOIRCSSNITGL 388
Db 387 QLFNSTWNTTIGPNTNG--TITLPCRIKOINRWQEVGKAMVAPPIRGOIRCSSNITGL 446
QY 389 LLTRDGGKETSNT-----TEIFRPGGDMRDNWSELYKYKVVVIEPLGVAPTAKRRVVQ 444
Db 447 LLTRDGGNSKNGSKNTEIFRPGGDMRDNWSELYKYKVVVIEPLGVAPTAKRRVVQ 506
QY 445 REKRAV-TLGAMFLGFLGAGSTWGAARSLTLTVQARQLLSGIVQQOQNLRLRAEQHLL 503
Db 507 REKRAVGTIGAMFLGFLGAGSTWGAARSLTLTVQARQLLSGIVQQOQNLRLRAEQHLL 566
QY 504 QLTWVGTK-LOARVLAVERYLKQQLLGIWCGSKLICTTAVPNWASWSNKSIDQWNNM 562
Db 567 QLTWVGTKLOARVLAVERYLKQQLLGIWCGSKLICTTAVPNWASWSNKSIDQWNNM 626
QY 563 TWMEWEIDNYNLTYLLEESQNEKEQELLELDKWASLWNNWFDISKWLWYIK 619
Db 627 TWMEWEIDNYNLTYLLEESQNEKEQELLELDKWASLWNNWFDISKWLWYIK 663
RESULT 4
ENV_HV1JR
ID _ENV HV1JR STANDARD; PRT; 848 AA.
AC P20871;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein gp160 precursor [Contains: Exterior membrane glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].
GN ENV
OS Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11688;
RN [1]
RA Koyanagi S., Chen I.S.Y.;
RL Submitted (DEC-1988) to the HIV data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M38429; AAB03749.1; -.
PDB; 1CE4; 18-MAR-99.
```

```
DR HIV; M38429; ENVJRCSF.
DR InterPro: IPR000328; Env GP41.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW Aids; Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 32
FT CHAIN 33 503
FT CHAIN 504 848
FT DISULFID 53 73
FT DISULFID 118 203
FT DISULFID 125 194
FT DISULFID 130 154
FT DISULFID 216 245
FT DISULFID 226 237
FT DISULFID 294 328
FT DISULFID 374 437
FT DISULFID 381 410
FT CARBOHYD 87 87
FT CARBOHYD 134 134
FT CARBOHYD 137 137
FT CARBOHYD 153 153
FT CARBOHYD 157 157
FT CARBOHYD 185 185
FT CARBOHYD 195 195
FT CARBOHYD 228 228
FT CARBOHYD 239 239
FT CARBOHYD 260 260
FT CARBOHYD 274 274
FT CARBOHYD 287 287
FT CARBOHYD 293 293
FT CARBOHYD 299 299
FT CARBOHYD 329 329
FT CARBOHYD 336 336
FT CARBOHYD 352 352
FT CARBOHYD 382 382
FT CARBOHYD 388 388
FT CARBOHYD 392 392
FT CARBOHYD 403 403
FT CARBOHYD 440 440
FT CARBOHYD 453 453
FT CARBOHYD 603 603
FT CARBOHYD 608 608
FT CARBOHYD 617 617
FT CARBOHYD 629 629
FT CARBOHYD 808 808
SQ SEQUENCE 848 AA; 96475 MW; 20767F51227EC3F3 CRC64;
Query Match 86.3%; Score 2870.5; DB 1; Length 848;
Best Local Similarity 83.7%; Pred. No. 8.6e-215;
Matches 544; Conservative 30; Mismatches 43; Indels 33; Gaps 5;
QY 1 SAVEKLVVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60
Db 28 SAAEQKLVVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 87
QY 61 VTENFNWKNMVEQMHEDIISLWDSKPCVKLTPLCVTLHCTNLKNAITNKNWEM 120
Db 88 VTEDFNWKNMVEQMHEDIISLWDSKPCVKLTPLCVTLHCTNLKNAITNKNWEM 146
QY 121 DRGEIKNCSFKVGAG-----KLINCVTSVITQACP 152
Db 147 ERGEIKNCSFNITTSIRSKVQKEYALFYKLDVVPIDNKNKTRLSNCTSVITQACP 206
QY 153 SFPEPIPIHYCAPAGFAILKNDKFPNGSGPCTNVSTVQCTHGIRPVVSTOLLNGSLAE 212
Db 207 SFPEPIPIHYCAPAGFAILKNNKTFNGKQCKNVSTVQCTHGIRPVVSTOLLNGSLAE 266
QY 213 GVVIRSENFTDNAKTIIVQLKESVEINCTRNPNNTTRSIHIGPGRFATGDIIGDIR 272
Db 267 KVVIRSDNFTDNAKTIIVQLNESVKINCTRNPNNTTRSIHIGPGRFATGDIIGDIR 326
```

QY 273 HCNISGEKWNNTLKQIVTKLOAQFGNKTIVFKOSSGDPBEIVMHSFNCGGEFFPCNSTQL 332  
Db 327 HCNISRAQWNNTLKQIVTKLOAQFGNKTIVFTSSGGDPBEIVMHSFNCGGEFFPCNSTQL 386  
QY 333 FNSWTNNTIGPNNTG--TTTLPCRKIQIINRQEVGKAMYPPIRQIRCSSNITGLLL 390  
Db 387 FNSWTNDTEKSSGTEGNDTTLPCRKIQIINRQEVGKAMYPPIRQIRCSSNITGLLL 446  
QY 391 TRDGGKEISNTTIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTAKRRVVQREKRAV 450  
Db 447 TRDGGKNESE-IEIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTAKRRVVQREKRAV 505  
QY 451 TLGAMFELGGAAGSTMGARSMTLTVOARQLLSGIQQQNNLRATEAQHLLQLTWGI 510  
Db 506 GIGALFLGFGAAGSTMGARSMTLTVOARQLLSGIQQQNNLRATEAQHLLQLTWGI 565  
QY 511 K-LQARVLAVERYLKQQLMGICGSKLCTTAVPNWASWSNKSLSDOITNNMTWMEWER 569  
Db 566 KQLOARVLAVERYLKQQLMGICGSKLCTTAVPNWASWSNKSLSDOITNNMTWMEWER 625  
QY 570 EIDNYTNLIYTLIEESQIQQEKNEQELLELDKWSLWNNFDSIKWLWYIK 619  
Db 626 EIDNYTNLIYTLIEESQIQQEKNEQELLELDKWSLWNNFDSIKWLWYIK 675  
RESULT 5  
ENV HV1A2  
ID ENV HV1A2 STANDARD; PRT; 855 AA.  
AC P03378;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85050453; PubMed=2578227;  
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,  
RA Stemplein M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,  
RA Levy J.A., Dina D., Lucin P.A.;  
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus  
(ARV-2).";  
RL Science 227:484-492 (1985).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; K02007; AAB59882.1; -.  
DR PIR; A03976; VCLJJA2.  
DR HIV; K02007; ENV\$SF2.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
FT SIGNAL 1 29  
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 208 BY SIMILARITY.  
FT DISULFID 125 199 BY SIMILARITY.  
FT DISULFID 130 155 BY SIMILARITY.  
FT DISULFID 221 250 BY SIMILARITY.

FT DISULFID 231 242 BY SIMILARITY.  
FT DISULFID 299 333 BY SIMILARITY.  
FT DISULFID 380 442 BY SIMILARITY.  
FT DISULFID 387 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;  
Query Match 85.8%; Score 2856; DB 1; Length 855;  
Best Local Similarity 82.3%; Pred. No. 1.2e-213;  
Matches 540; Conservative 32; Mismatches 46; Indels 38; Gaps 6;  
QY 1 SAVEKLWTVVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLEN 60  
Db 28 SATEKLWTVVYGVVPVWKEATTLFCASDARAYDTEVHNWATHACVPTDNPQEVVLGN 87  
QY 61 VTENFNWNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLHCNLKNAINTKSSNWKEM 120  
Db 88 VTENFNWNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLHCNLKNAINTKSSNWKEE 147  
QY 121 DRGEIKNCSFKVAG-----KLINCNTSVITQA 148  
Db 148 IKGEIKNCSFNITTSIRDKIQKENALFRNLDVVPIDNASTTTNTYNYRLIHCNRSVITQA 207  
QY 149 CPKVSPEPIPIHYCAPAGPAILKNDKPKNGSGPCTNVSTVQCTHGIRVPVSTQLLINGS 208  
Db 208 CPKVSPEPIPIHYCTPAGPAILKCNKTFNGKGPCTNVSTVQCTHGIRPIVSTQLLINGS 267  
QY 209 LAEEGVVIRSENFNTONAKTIIIVOLKESVEINCTRPNNNTKRSITICGPAFYATGDIIGD 268  
Db 268 LAEEVVIRSDNFTNNAKTIIVOLNESVAINCTRPNNNTKRSIYIGPGAFTHTGRIIGD 327  
QY 269 IROAHCNISGEKWNNTLKQIVTKLOAQFG-NKTIVFKOSSGGDPBEIVMHSFNCGGEFFYC 327  
Db 328 IRKAHCNISRAQWNNTLEQIVTKLREQFGNKTIVFNQSSGGDPBEIVMHSFNCGGEFFYC 387  
QY 328 NSTQLFNSWT--NNTIGPNNTNGTITLPCRKIQIINRQEVGKAMYPPIRQIRCSSNI 395  
Db 388 NNTQLFNSWTNRLNHTEGTKG-NDTILPCRKIQIINRQEVGKAMYPPIRQIRCSSNI 446  
QY 386 TGLLLTRDGGKEISNTTIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTAKRRVVOR 445  
Db 447 TGLLLTRDGGTNTVNDTEVFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTAKRRVVOR 506

QY 446 EKRAV-TLGAFLGFLGAAGSTWARSLLTVQARQLLSGIQQNNLLRAEAQOHLQ 504  
 Db 507 EKRAVIGVGNFGLGAGSTWAGVSLTLTVQARQLLSGIQQNNLLRAEAQOHLQ 566  
 QY 505 LTVWGK-IOARVLAVERYKQDLGIGWCSGLICTTAVPNWASWNSKSLDQIWNMT 563  
 Db 567 LTVWGKIQARVLAVERYLQDLGIGWCSGLICTTAVPNWASWNSKSLDQIWNMT 626  
 QY 564 WMEWERIDYNTLIVLTLESQOQEKNEOELELDKWSLWNNWEDISKWLYIK 619  
 Db 627 WMQWERIDYNTTITLTLESQOQEKNEOELELDKWSLWNNWFSITNWLWYIK 682

RESULT 6

ENV\_HV1B1  
 ID ENV\_HV1B1 STANDARD; PRT; 856 AA.  
 AC P03375;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11678;  
 RN [1].  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8511123; PubMed=2578615;  
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
 RA Josephs S.F., Doran E.R., Rafaleki J.A., Whitehorn E.A.,  
 RA Baumaister K., Ivanoff L., Pettaway S.R. Jr, Pearson M.L.,  
 RA Lautenberger J.A., Papas T.S., Ghirayeb J., Chang N.T., Gallo R.C.,  
 RA Wong-Staal F.;  
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
 RL Nature 313:277-284 (1985).  
 RN [2].  
 RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.  
 RX MEDLINE=90285159; PubMed=2355006;  
 RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,  
 RA Gregory T.J.;  
 RT "Assignment of intrachain disulfide bonds and characterization of  
 RT potential glycosylation sites of the type 1 recombinant human  
 RT immunodeficiency virus envelope glycoprotein (gp120) expressed in  
 RT Chinese hamster ovary cells.";  
 RL J. Biol. Chem. 265:10373-10382 (1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M15654; AAA4205.1; --  
 DR PIR: A03973; VCLJH3.  
 DR HIV: M15654; ENV\$BH102.  
 DR InterPro: IPR000328; Env GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239

FT DISULFID 296 331  
 FT DISULFID 378 445  
 FT DISULFID 385 418  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97224 MW; 0BFBI18931BB27 CRC64;  
 Query Match 85.7%; Score 2851; DB 1; Length 856;  
 Best Local Similarity 83.4%; Pred. No. 2.9e-213;  
 Matches 548; Conservative 32; Mismatches 37; Indels 40; Gaps 10;  
 QY 1 SAVEKLWVTYVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60  
 Db 29 SATEKLWVTYVYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 88  
 QY 61 VTENFNWNNWNNVQNHEDIISLWDSLKPCVKLTPLCVTLHCTNLKNAITNTSSNWK 119  
 Db 89 VTENFNWNNWNNVQNHEDIISLWDSLKPCVKLTPLCVTLHCTNLKNAITNTSSNWK 148  
 QY 120 MDRGEIKNGSFYKGA---GK-----LINCNTSVITQACPKV 152  
 Db 149 MEKGEIKNGSFNISTSGIRKGVQKEYAFYKLDIIPIDNDTTSYTLTSCNTSVITQACPKV 208  
 QY 153 SPEPIPIHYCAPAGFAILKCNKKFNGSGPCTNVTSTVQCTHGRPVVSTOLLNGSLAE 212  
 Db 209 SPEPIPIHYCAPAGFAILKCNKNKTGCTNVTSTVQCTHGRPVVSTOLLNGSLAE 268  
 QY 213 GVVISRENTDNAKTIIVOLKESVEINCTRPNNTRKSIITI--GPGRAFVATGDIGDIR 270  
 Db 269 EVVISRANFTDNAKTIIVOLNOSVEINCTRPNNTRKSIIRIORGPGRAFVATGK--IGNMR 327  
 QY 271 QAHCNISGEKNWNTLQIVTKLQAFQF-NKTIIVFKOSSGGDPPIVMSFNCGGEFFYCNS 329  
 Db 328 QAHCNISRAKNWNTLQKIDSKLEQFGNNKTIIFKSSGGDPPIVTHSFCGGEFFYCNS 387  
 QY 330 TQLFNSTWNN----TIGPNTNG--TITLPCRKIQIINRWQEVGKAWYAPPINGQIRCSS 383  
 Db 388 TQLFNSTWNTSTWSTKSGNNTGSDTITLPCRKIQIINRWQEVGKAWYAPPISQIRCSS 447  
 QY 384 NITGLLLTRDGGKEISNTTTEIFRPGGGDMRDNWSELYKYVKVIEPLGVAPTKAKRRV 443  
 Db 448 NITGLLLTRDGGNS--NNSEIFRPGGGDMRDNWSELYKYVKVIEPLGVAPTKAKRRV 506  
 QY 444 QREKRAVTLGAMPLGLGAAGSTWARSLLTVQARQLLSGIQQNNLLRAEAQOHL 503







AC P03376;  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11700;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111157; PubMed=2982104;  
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Laaky L.A.,  
RT "Nucleic acid structure and expression of the human  
RL AIDS/lymphadenopathy retrovirus";  
RL Nature 313:450-458 (1985).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; K02083; AAB59873.1; --  
DR EMBL; X01762; CAA25903.1; ALT\_SEQ.  
DR PIR; A03974; VCLJVL.  
DR HIV; K02083; ENVSPV22.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
KW Signal.  
FT SIGNAL 1 30  
FT CHAIN 31 511  
FT CHAIN 512 856  
FT DISULFID 54 74  
FT DISULFID 119 205  
FT DISULFID 126 196  
FT DISULFID 131 157  
FT DISULFID 218 247  
FT DISULFID 228 239  
FT DISULFID 296 331  
FT DISULFID 378 445  
FT DISULFID 385 418  
FT CARBOHYD 88 88  
FT CARBOHYD 136 136  
FT CARBOHYD 141 141  
FT CARBOHYD 156 156  
FT CARBOHYD 160 160  
FT CARBOHYD 186 186  
FT CARBOHYD 197 197  
FT CARBOHYD 230 230  
FT CARBOHYD 234 234  
FT CARBOHYD 241 241  
FT CARBOHYD 262 262  
FT CARBOHYD 276 276  
FT CARBOHYD 289 289  
FT CARBOHYD 295 295  
FT CARBOHYD 301 301  
FT CARBOHYD 332 332  
FT CARBOHYD 339 339  
FT CARBOHYD 356 356  
FT CARBOHYD 386 386  
FT CARBOHYD 392 392  
FT CARBOHYD 397 397  
FT CARBOHYD 406 406  
FT CARBOHYD 448 448  
FT CARBOHYD 463 463

FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 856 AA; 97339 MW; 5FCDB1DC3C1209B3 CRC64;  
Query Match 85.3%; Score 2838; DB 1; Length 856;  
Best Local Similarity 83.0%; Pred. No. 2.9e-212;  
Matches 545; Conservative 33; Mismatches 39; Indels 40; Gaps 10;  
QY 1 SAVEKLWTVYVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60  
DB 29 SATEKLWTVYVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVLYN 88  
QY 61 VTENFMNKNWVQWQHEDIISLWDSLXPCVKLTPLCVTLHCTNLKATNTKSSNWE- 119  
DB 89 VTENFMNKNWVQWQHEDIISLWDSLXPCVKLTPLCVTLHCTNLKATNTNTSSGMI 148  
QY 120 MDRGEIKNCSEKVA--GK-----LINCNTSVITQCPKV 152  
DB 149 MEKGEIKNCSEKVA--GK-----LINCNTSVITQCPKV 208  
QY 153 SPEPIPIHYCAPAGFAILKCNCKFKNGSGPCNTNVSTVQCTHGIRPVVSTQLLNGSLAE 212  
DB 209 SPEPIPIHYCAPAGFAILKCNCKFKNGSGPCNTNVSTVQCTHGIRPVVSTQLLNGSLAE 268  
QY 213 GVVISSEFTDNAKTIIVOLKESVEINCTRPNNTRKSGITI--GPGRAFATGDIIGDIR 270  
DB 269 EVVIRSANFTDNAKTIIVOLKESVEINCTRPNNTRKSGIRIQRGPGRAFVTIGK-IGNMR 327  
QY 271 QAHCNISGEKWNNTLKQIVTKLQAFQ-NKTIIVKQSSGSDPEIVMHSFNCGGEFFYCN 329  
DB 328 QAHCNISRAKNNTLKQIDSKLEQFGNNKTIIFKQSSGSDPEIVMHSFNCGGEFFYCN 387  
QY 330 TQLFNSTWNN---TIGPNNTNG--TITLPCRIKQIINRWQEVGKAMYPPIRGQIRCSS 383  
DB 388 TQLFNSTWNN---TIGPNNTNG--TITLPCRIKQIINRWQEVGKAMYPPIRGQIRCSS 447  
QY 384 NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKARRVV 443  
DB 448 NITGLLLTRDGGNN--NNESEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKARRVV 506  
QY 444 QREKRAVTIGAMFLGFLGAGSTMGARSLTLTVQARQLLSGIVQOQNLLRAEAQHLL 503  
DB 507 QREKRAVGIGALFLGFLGAGSTMGARSLTLTVQARQLLSGIVQOQNLLRAEAQHLL 566  
QY 504 QLTVMGK-LQARVLAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDQIWNMM 562  
DB 567 QLTVMGKIQARILAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDQIWNMM 626  
QY 563 TMWEVEREIDNTNLIYTLIESQKQKNEQELLELDKWLWNSFDFISKMLWYIK 619  
DB 627 TMWEVEREIDNTNLIYTLIESQKQKNEQELLELDKWLWNSFDFISKMLWYIK 683  
RESULT 10  
ENV\_HV1H2  
ID ENV\_HV1H2 STANDARD; PRT; 856 AA.  
AC P04578; O09779;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11706;  
RN [1]  
RP SEQUENCE FROM N.A.



SEQUENCE FROM N.A.  
 MEDLINE=85111123; PubMed=2578615;  
 RX Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
 RA Baumanister K., Ivanoff L., Pettaway S.R. Jr., Pearson M.L.,  
 RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,  
 RA Wong-Staal F.;  
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
 RL Nature 313:277-284 (1985).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; KO2011; AAA44661.1; -;  
 DR PDB; 1DDH; 13-JAN-99.  
 DR PDB; 1HHG; 31-OCT-93.  
 DR PDB; 1Q03; 02-JAN-00.  
 DR PDB; 1S2T; 24-DEC-97.  
 DR HIV; KO2011; ENV\$B88.  
 DR GlycoSuiteDB; P04582; -;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 30  
 FT CHAIN 31 506 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 507 851 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74 BY SIMILARITY.  
 FT DISULFID 119 205 BY SIMILARITY.  
 FT DISULFID 126 196 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 218 247 BY SIMILARITY.  
 FT DISULFID 228 239 BY SIMILARITY.  
 FT DISULFID 296 331 BY SIMILARITY.  
 FT DISULFID 378 440 BY SIMILARITY.  
 FT DISULFID 385 413 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 606 606 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 632 632 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 811 811 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 85.0%; Score 2829.5; DB 1; Length 851;  
 Best Local Similarity 83.0%; Pred. No. 1.3e-211;  
 Matches 542; Conservative 33; Mismatches 41; Indels 37; Gaps 9;  
 QY 1 SAVELKLVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNQEIVLEN 60  
 DB 29 SATEKLWTVYFVGPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNQEIVLVN 88  
 QY 61 VTENFNWNNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLHCTNKNATNTKSSNWE- 119  
 DB 89 VTENFNWNNMVEQMHEDIISLWDSQSLKPCVKLTPLCVSLKCTDLKNDNTNSSGRMI 148  
 QY 120 MRGEIKNSFKVGAOK-----LINCSVTITQACPVK 152  
 DB 149 MEKGEIKNSFNISTSKRGVKEYAFFVKLDIIPDNDTTSYTLTSCNTSVITQACPVK 208  
 QY 153 SPEPIHYCAPAGFAILKNDKFNKSGPCNTVSTVQCTHGRPVVSTOLLNGSLAE 212  
 DB 209 SPEPIHYCAPAGFAILKNNKNTFNGTGTNTVSTVQCTHGRPVVSTOLLNGSLAE 268  
 QY 213 GVVIRSEFTDNAKTIIVOLKESVEINCTPNNNTRKSIITII--GPGRFAVATGDIIDIR 270  
 DB 269 EVVIRSEFTDNAKTIIVQDTSVEINCTPNNNTRKIRIQRGPGRAFTVIGK-IGNMR 327  
 QY 271 QAHCNISGEKNWNTLKOIVTKLQAFQ-NKTIIVFKOSSGGDPEIVMHSFNCGGEFFYCNS 329  
 DB 328 QAHCNISRAKNWNTLKOIDSKLREQFGNNTIIFKOSSGGDPEIVTHSFNCGGEFFYCNS 387  
 QY 330 TQLFNSWNTTGPNNNG--TITLPCRIRKQIINRQWCKYKAPPIRQICSSNITG 387  
 DB 388 TQLFNSW-STKGSNNTEGSDTITLPCRIRKQIINRQWCKYKAPPIRQICSSNITG 446  
 QY 388 LLLTRDGGKBISSNTTIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTAKKRVVOREK 447  
 DB 447 LLLTRDGGNS-NNESEIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTAKKRVVOREK 505  
 QY 448 RAVTLGAMFLGFGAAGSTMGARSLLTVQARQLLSGIVQOQNNLLRAIEAQHLLQLTV 507  
 DB 506 RAVGIGALFLGFGAAGSTMGAAASMTLTQARQLLSGIVQOQNNLLRAIEAQHLLQLTV 565  
 QY 508 WGIK-LQARVLAVERYLKQQLLGIWGCCKLCTTAVPMNWSNKSLSQIWNNTTME 566  
 DB 566 WGIKQARILAVERYLKDQLLGIWGCCKLCTTAVPMNWSNKSLSQIWNNTTME 625  
 QY 567 WEREINYNTLIYTLIEESQOQKNEQLLELDKWLASLWNPFDISKWLWYIK 619  
 DB 626 WDREINNYSLHSIEESQOQKNEQLLELDKWLASLWNPFDISKWLWYIK 678

## RESULT 12

ENV\_HV1BN STANDARD; PRT; 852 AA.  
 AC P12488;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (HIV isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11693;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89085613; PubMed=2789516;  
 RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,  
 RA Dandekar S.;  
 RT "Biological and molecular characterization of human immunodeficiency virus (HIV-1BR) from the brain of a patient with progressive dementia.";  
 RT Virology 168:79-89 (1989).  
 CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS











FT	CARBOHYD	88	88	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	136	136	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	141	141	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	156	156	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	160	160	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	186	186	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	197	197	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	230	230	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	234	234	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	241	241	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	262	262	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	276	276	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	289	289	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	295	295	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	301	301	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	332	332	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	339	339	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	356	356	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	386	386	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	392	392	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	397	397	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	406	406	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	448	448	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	463	463	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	611	611	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	616	616	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	624	624	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	637	637	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	674	674	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	750	750	N-LINKED	(GLCNAC. .)	(POTENTIAL)
FT	CARBOHYD	816	816	N-LINKED	(GLCNAC. .)	(POTENTIAL)
SQ	SEQUENCE	856 AA; 97188 MW; 3373C68B84C1AFC CRC64;				
Query Match: 84.4%; Score 2808; DB 1; Length 856;						
Best Local Similarity 82.3%; Pred. No. 6.2e-210;						
Matches 541; Conservative 33; Mismatches 43; Indels 40; Gaps 10;						
Qy	1	SAVEKLVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN	60			
Db	29	SATEKLVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHAGVPTDPNPOEVVLVN	88			
Qy	61	VTFNFMWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNLGNATNTKSSNWKE-	119			
Db	89	VTFNFMWKNMVEQMHEDIISLDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSGRMI	148			
Qy	120	MDRGEIKNCSFKVGA---GK-----LINCNTSVITQACP	152			
Db	149	MEKGEIKNCSFNISTSRGKQKEYAFYKLDIIPIDNDTTSYTLTSCNTSVITQACP	208			
Qy	153	SFEPPIHYCAPAGFAILKCNCKFNGSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAE	212			
Db	209	SFEPPIHYCAPAGFAILKCNCKTFNGTGVCTNVSTVQCTHGIRPVVSTQLLNGSLAE	268			
Qy	213	GVVIRSENFTDNAKTIIVQLKESVEINCTRPNNTKSIIT--GPGRAFYATGDIIGDIR	270			
Db	269	EVVIRSVNFTDNAKTIIVQLTVEINCTRPNNTKRIQIRGPGRAFVTIGK-IGNMR	327			
Qy	271	QAHCNISGEKNNTLKOIVTKLQAFQ-NKTIIVFKQSSGGDPETVMSFNCGGEFFYCN	329			
Db	328	QAHCNISRAKNATLKQIAKLRQFGNNKTIIPKQSSGGDPETVMSFNCGGEFFYCN	387			
Qy	330	TQLFNSWN--TIGPNNTNG--TITLPCRKQIINRWQEVGKAMYAPPPIRQIRCSS	383			
Db	388	TQLFNSWTFNSTWTEGSNTEGSDTITLPCRKQFINMWQEVGKAMYAPPISQIRCSS	447			
Qy	384	NITGLLLTRDGGKEISNTTEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKKRRV	443			
Db	448	NITGLLLTRDGGNN--NNGSEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKKRRV	506			
Qy	444	QREKRAVTLGAMFLGFGAGSTWGAESLTI-TVOAROLLGSIVOOQNLLRATEAQOHL	503			
Db	507	QREKRAVGI GALFUGFGAGSTWGAASMTLT-TVOAROLLGSIVOOQNLLRATEAQOHL	566			

Qy	504	QLTVWGIK-LOARVLAVERYLKDOQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNMM	562
Db	567	QLTVWGIKQLQARILAVERYLKDOQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNHT	626
Qy	563	TWMEWREIDNYTNLIIVTLTEESQOQEKNEQELLELDKWASLWNWFDISKWLWYIK	619
Db	627	TWMEWREINNYTSLIHSLEESQOQEKNEQELLELDKWASLWNWFNITNLWYIK	683

Search completed: February 25, 2004, 14:17:01  
Job time : 12.7652 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 14:12:26 ; Search time 39.1462 Seconds  
(without alignments)  
4989.132 Million cell updates/sec

Title: US-09-891-609A-4  
Perfect score: 3327  
Sequence: 1 SAVEKLVTVYGVVPVWKEA.....DKWASLWNWPFDISKWLWYIK 619

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3176	95.5	847	15	Q8ar21 human immun
2	2978.5	89.5	854	15	Q7SVL3 human immun
3	2971.5	89.3	850	15	Q7SVL4 human immun
4	2968.5	89.2	860	15	Q7SVL7 human immun
5	2967.5	89.2	852	15	Q7SVL5 human immun
6	2941	88.4	853	15	Q7SVL6 human immun
7	2925	87.9	843	15	Q70150 human immun
8	2924.5	87.9	854	15	Q40222 human immun
9	2914.5	87.6	846	15	Q9PXW7 human immun
10	2906	87.3	847	15	Q75760 human immun
11	2900.5	87.2	852	15	Q41883 human immun
12	2895.5	87.0	854	15	Q78225 human immun
13	2895	87.0	851	15	O56110 human immun
14	2894	87.0	854	15	O92762 human immun
15	2893	87.0	853	15	O56108 human immun
16	2892	86.9	853	15	Q03811 human immun

17	2890	86.9	853	15	Q7ZJF5 human immun
18	2889.5	86.9	852	15	O92761 human immun
19	2887.5	86.8	856	15	Q77694 human immun
20	2886.5	86.8	858	15	O7SVL1 human immun
21	2886	86.7	845	15	O70679 human immun
22	2883.5	86.7	850	15	Q74812 human immun
23	2883.5	86.7	860	15	Q7ZJDF7 human immun
24	2882.5	86.6	864	15	Q8UL53 human immun
25	2881.5	86.6	856	15	Q72993 human immun
26	2878	86.5	859	15	Q7SUT0 human immun
27	2876.5	86.5	855	15	Q03805 human immun
28	2875	86.4	855	15	Q8Q367 human immun
29	2874.5	86.4	868	15	Q9WJW5 human immun
30	2874	86.4	859	15	Q7SVL2 human immun
31	2871	86.3	849	15	Q77368 human immun
32	2870.5	86.3	867	15	Q7ZC00 human immun
33	2869	86.2	861	15	Q9IUY9 human immun
34	2869	86.2	863	15	Q77989 human immun
35	2868	86.2	859	15	Q7SUS9 human immun
36	2867.5	86.2	852	15	O41885 human immun
37	2867	86.2	853	15	Q7SUR2 human immun
38	2867	86.2	853	15	Q7SIK0 human immun
39	2865	86.1	861	15	O9WJTV5 human immun
40	2865	86.1	863	15	O42031 human immun
41	2864.5	86.1	854	15	Q7SUQ7 human immun
42	2864	86.1	861	15	O56109 human immun
43	2863.5	86.1	846	15	Q8UL63 human immun
44	2863	86.1	853	15	Q7SUR3 human immun
45	2862.5	86.0	854	15	Q7SUQ9 human immun

ALIGNMENTS

RESULT 1

Q8AR21	ID	Q8AR21	PRELIMINARY;	PRT;	847 AA.
AC	Q8AR21;				
DT	01-MAR-2003 (TrEMBLrel. 23, Created)				
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)				
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)				
DE	Envelope glycoprotein.				
GN	ENV.				
OS	Human immunodeficiency virus 1.				
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxID=11676;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SHIVSF162P3.2;				
RA	Gao F.;				
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF536757; AAN05642.1; -				
DR	GO; GO:0019031; C:virial envelope; IEA.				
DR	GO; GO:0005198; F:structural molecule activity; IEA.				
DR	InterPro; IPR000328; Env_GP41.				
DR	InterPro; IPR000777; GP120.				
DR	Pfam; PF00516; GP120; 1.				
DR	Pfam; PF00517; GP41; 1.				
SQ	SEQUENCE 847 AA; 95879 MW; 9CF863B97B3BB54C CRC64;				

Query Match 95.5%; Score 3176; DB 15; Length 847;  
Best Local Similarity 92.1%; Pred. No. 2e-250;  
Matches 596; Conservative 7; Mismatches 16; Indels 28; Gaps 2;

Qy	1	SAVEKLVTVYGVVPVWKEATTTLFCSADAKAYDTEVHNWATHACVPTDPNPOEIVLEN 60
Db	28	SAVEKLVTVYGVPAWKEATTTLFCSADAKAYDTEVHNWATHACVPTDPNPOEIVLEN 87
Qy	61	VTFNFFNKKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNATNTKSSNNKEM 120
Db	88	VTFNFFNKKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNATNTKSSNNKEM 147
Qy	121	DRGEIKNCSPKVGAG-----KLINCNTSVITQACPKVS 153

```
Db 148 NRGEIKNCSFNVTTISGNKMQKEYALFYRLDVVPIDNDNTSYNLNCNTSVITQACPKVS 207
:||||| |
Qy 154 PEPIPIHYCAPAGFAILKCNDDKFKNGSGPCNTVSTVQCTHGIRPVVSTQLLNGSLAEG 213
|||||
Db 208 PEPIPIHYCAPAGFAILKCNDDKFKNGSGPCINVTQCTHGIRPVVSTQLLNGSLAEG 267
|||||
Qy 214 VVIRENFDTNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGAFYATGDIIGDIRQAH 273
|||||
Db 268 VVIRENFDTNAKTIIVQLKESVEINCTRPNNNTRKSIIPGPKAFYATGDIIGDIRQAH 327
|||||
Qy 274 CNISGEKWNNTLKQIVTKLQAFQGNKTIIVFKQSGGDPPEIVMHSFNCGGEFFYCNSTQLF 333
|||||
Db 328 CNISGEKWNNTLKQIVTKLQAFQGNKTIIVFKQSGGDPPEIVMHSFNCGGEFFYCNSTQLF 387
|||||
Qy 334 NSTWNNTGPNNTNGTITLPCRICKQIINRWQEVGKAMYAPPPIRGQIRCSSNITGLLTRD 393
|||||
Db 388 NSTWNNTGPNNTNGTITLPCRICKQIINRWQEVGKAMYAPPPIRGQIRCSSNITGLLTRD 447
|||||
Qy 394 GKGKISNTTEIFRPGGDMRDNWRSSELYKYKVKIEPLGVAPTAKKRRVVOREKRAVTLG 453
|||||
Db 448 GGREVGNTTEIFRPGGDMRDNWRSSELYKYKVKIEPLGVAPTAKKRRVVOREKRAVTLG 507
|||||
Qy 454 AMFLGFLGAAGSTWGAARSLTITVOARQLLSGIVQOQNLLRAIEAQHLLQLTVMGIK-L 512
|||||
Db 508 AVFLGFLGAAGSTWGAARSLTITVOARQLLSGIVQOQNLLRAIEAQHLLQLTVMGIKQL 567
|||||
Qy 513 QARVLAVERYLKDQQLLGWCSGKLICTTAVPNWNASWSKSLDQIWNNTMTMWEEREID 572
|||||
Db 568 QARVLAVERYLKDQQLLGWCSGKLICTTAVPNWNASWSKSLDQIWNNTMTMWEEREIG 627
|||||
Qy 573 NYTNLIYTLIEESQOQKNEQELLELDKWASLWNPEDISKWLWYIK 619
|||||
Db 628 NYTNLIYTLIEESQOQKNEQELLELDKWASLWNPEDISKWLWYIK 674
|||||
```

## RESULT 2

```
Q7SVL3 PRELIMINARY; PRT; 854 AA.
ID Q7SVL3
AC Q7SVL3;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=81NY1;
RX MEDLINE=22628496; PubMed=12743293;
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairoj A.S.,
RA Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;
RT "U.S. Human Immunodeficiency Virus Type 1 Epidemic; Date of Origin,
RT Population History, and Characterization of Early Strains.";
RL J. Virol. 77:6359-6366(2003).
DR EMBL; AY247222; AAP37150.1; -.
KW Envelope protein.
SQ SEQUENCE 854 AA; 97048 MW; 36379231PF3AF12C CRC64;
```

Query Match 89.5%; Score 2978.5; DB 15; Length 854;  
Best Local Similarity 85.9%; Pred. No. 2.8e-234;  
Matches 562; Conservative 21; Mismatches 36; Indels 35; Gaps 5;

```
Qy 1 SAVEKLVTVYVGVVPWKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQBIIVEN 60
|||
Db 28 SAAEQKLVTVYVGVVPWKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQBIIVEN 87
|||
Qy 61 VTENFNWKNWNNVQOMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNATNTKSNWKEM 120
|||||
Db 88 VTENFNWKNWNNVQOMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNATNTKSNWKEM 147
|||||
```

```
Qy 121 DRGEIKNCSFKVGAG-----KLNCNNTSVITQACPK 151
:|||||
Db 148 ESGEIKNCSFNTTTSIRDKVQKEYALFYKLDVVPINDNTTTSYRLNCNNTSVITQACPK 207
|||||
Qy 152 VSPEPIPIHYCAPAGFAILKCNDDKFKNGSGPCNTVSTVQCTHGIRPVVSTQLLNGSLAE 211
|||||
Db 208 VTPEPIPIHYCTPAGEFAILKCNDDKFKNGSGPCNTVSTVQCTHGIRPVVSTQLLNGSLAE 267
|||||
Qy 212 EGVVIRENFDTNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGAFYATGDIIGDIRQ 271
|||||
Db 268 EGVVIRENFDTNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGAFYATGDIIGDIRQ 327
|||||
Qy 272 AHNCNISEKWNNTLKQIVTKLQAFQGNKTIIVFKQSGGDPPEIVMHSFNCGGEFFYCNST 330
|||||
Db 328 AHNCNLSAKKNDTLKQIVTKLREQFGNKNKTIIVFNQSGGDPPEIVMHSFNCGGEFFYCDST 387
|||||
Qy 331 QLFNSTW---NNTIGPNNTNGTITLPCRICKQIINRWQEVGKAMYAPPPIRGQIRCSSNITG 387
|||||
Db 388 QLFNSTWENDNNIRGNSNSTQENITLPCRICKQIINRWQEVGKAMYAPPPIRGQIRCSSNITG 447
|||||
Qy 388 LLLTRDGGKESISNTTEIFRPGGDMRDNWRSSELYKYKVKIEPLGVAPTAKKRRVVOREK 447
|||||
Db 448 LLLTRDGGKDENGTTEIFRPGGDMRDNWRSSELYKYKVKIEPLGVAPTAKKRRVVOREK 507
|||||
Qy 448 RAV-TLGAFLGFLGAAGSTWGAARSLTITVOARQLLSGIVQOQNLLRAIEAQHLLQLT 506
|||||
Db 508 RAVGTIGAMFLGFLGAAGSTWGAARSLTITVOARQLLSGIVQOQNLLRAIEAQHLLQLT 567
|||||
Qy 507 VMGIK-LOARVLAVERYLKDQQLLGWCSGKLICTTAVPNWNASWSKSLDQIWNNTMTM 565
|||||
Db 568 VMGIKQLQARVLAVERYLKDQQLLGWCSGKLICTTAVPNWNASWSKSLDQIWNNTMTM 627
|||||
Qy 566 EWEREIDNYTNLIYTLIEESQOQKNEQELLELDKWASLWNPEDISKWLWYIK 619
|||||
Db 628 EWEREIDNYTNLIYTLIEESQOQKNEQELLELDKWASLWNPEDISKWLWYIK 681
|||||
```

## RESULT 3

```
Q7SVL4 PRELIMINARY; PRT; 850 AA.
ID Q7SVL4
AC Q7SVL4;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=81NJ;
RX MEDLINE=22628496; PubMed=12743293;
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairoj A.S.,
RA Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;
RT "U.S. Human Immunodeficiency Virus Type 1 Epidemic; Date of Origin,
RT Population History, and Characterization of Early Strains.";
RL J. Virol. 77:6359-6366(2003).
DR EMBL; AY247221; AAP37149.1; -.
KW Envelope protein.
SQ SEQUENCE 850 AA; 96583 MW; 62ED5F6AB033D20B CRC64;
```

Query Match 89.3%; Score 2971.5; DB 15; Length 850;  
Best Local Similarity 86.2%; Pred. No. 1e-233;  
Matches 561; Conservative 24; Mismatches 33; Indels 33; Gaps 4;

```
Qy 1 SAVEKLVTVYVGVVPWKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQBIIVEN 60
|||
Db 28 SAAEQKLVTVYVGVVPWKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQBIIVEN 87
|||
Qy 61 VTENFNWKNWNNVQOMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNATNTKSNWKEM 120
|||||
Db 88 VTENFNWKNWNNVQOMHEDIISLWQSLKPCVKLTPLCVTLHCTNLKNATNTKSNWKEM 147
|||||
```



```
Db      88 VTFNNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLGNTNTTSSSGGGM 147
Qy      121 DRGEIKNCSPKVGAG-----KLNCHNTSVITQACPKV 153
Db      148 ERGEIKNCSEFNITTSIRDKVQKEYALLYKLDIVPIDNDNTSYRLISCNSTSVITQACPKV 207
Qy      154 FEPIPIHYCAPAGFAILKNDKFKNGSGPCNTNVSTVQCTHGIRPVVSTQLLLNGSLAEEG 213
Db      208 FEPIPIHYCAPAGFAILKCDKFKNGSGPCNTNVSTVQCTHGIRPVVSTQLLLNGSLAEE 267
Qy      214 VVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITGPGRAFATGDIIGDIRQAH 273
Db      268 VVIRSDNFDTNAKTIIVQLKESVEINCTRPNNTRKSIHIGPGRAFATGDIIGDIRQAH 327
Qy      274 CNTISGEKNNNTLKQIVTKLQAFQNKTIIVFKQSSGGDPEIVMHSFNCGGEPFYCNSTOLF 333
Db      328 CNTISKAKWNTTLKQIVTKLQAFQNKTIIVFNQSSGGDPEIVMHSFNCGGEPFYCNSTOLF 387
Qy      334 N-----STWNNTIGPNNNTGTTILPCRKQIINRMQEVGKAMYAPPIRGQIRCSSNITG 387
Db      388 NSTWNASTWNTDTEGSSNTEGTTILPCRKQIINRMQEVGKAMYAPPIRGQIRCSSNITG 447
Qy      388 LLTLTRDGGKEISNTEIFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKARRVVQREK 447
Db      448 LILTRDGGSN-ENDETEIFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKARRVVQREK 506
Qy      448 RAVTLGAMFGLGFLGAAGSTWGARSLLTVQARQLLSGIVQOQNNLLRAIEAQHLLQITV 507
Db      507 RAVGIGAVFLGFLGAAGSTWGAAASMTLTVQARQLLSGIVQOQNNLLRAIEAQHLLQITV 566
Qy      508 WGIK-LOARVLAVERYLKQDOLLGWCSCGLICTTAVPWNASNSKSLDOINNNMTWME 566
Db      567 WGIKQLQARVLAVERYLKQDOLLGWCSCGLICTTAVPWNASNSKSLDOINNNMTWME 626
Qy      567 WEREIDNNTLIYTLIEESQKNEQELLEDKWSLNNWFDISKWLWYIK 619
Db      627 WDRINNTSYLIYTLIEESQKNEQELLEDKWSLNNWFDITKWLWYIK 679

RESULT 6
Q7SVL6
ID Q7SVL6 PRELIMINARY; PRT; 853 AA.
AC Q7SVL6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=81CA2;
RX MEDLINE=22628496; PubMed=12743293;
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairoj A.S.,
RA Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;
RA "U.S. Human Immunodeficiency Virus Type 1 Epidemic: Date of Origin,
RT Population History, and Characterization of Early Strains.";
RL J. Virol. 77:6359-6366(2003).
DR EMBL; AY247219; AAP37147.1; -.
KW Envelope protein.
SQ SEQUENCE 853 AA; 96545 MW; E3B2830A1261E237 CRC64;
```

```
Query Match      88.4%; Score 2941; DB 15; Length 853;
Best Local Similarity 84.8%; Pred. No. 3.2e-231;
Matches 556; Conservative 30; Mismatches 30; Indels 40; Gaps 5;

Qy      1 SAVEKLWTVYGVDPVWKEATTTILFCASDAKAYDTEVHNVWATHACVPTDPNQEIVLEN 60
Db      28 SATEKLWTVYGVDPVWKEATTTILFCASDAKAYDTEVHNVWATHACVPTDPNQEIVLGN 87
```

```
Qy      61 VTFNNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTLNKATNTKSSNWKEM 120
Db      88 VTFNNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLRATNTTSSSGGKM 147
Qy      121 DRGEIKNCSPKVGAG-----KLNCHNTSVITQACPK 151
Db      148 EGGEIKNCSEFNITTSIRDKMKQKEYALFYKLDVVVPIDNDNTSYRLISCNSTSVITQACPK 207
Qy      152 VSPEPIPIHYCAPAGFAILKNDKFKNGSGPCNTNVSTVQCTHGIRPVVSTQLLLNGSLAE 211
Db      208 VSPEPIPIHYCAPAGFAILKNDKFKNGSGPCNTNVSTVQCTHGIRPVVSTQLLLNGSLAE 267
Qy      212 EGVVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITGPGRAFATGDIIGDIRQ 271
Db      268 EGVVIRSDNFDTNAKTIIVQLKESVEINCTRPNNTRKSIHIGPGRAFATGDIIGDIRQ 327
Qy      272 AHCHNISEGKNNNTLKQIVTKLQAFQNKTIIVFKQSSGGDPEIVMHSFNCGGEPFYCNSTQ 331
Db      328 AHCHNLSITKWNNTLKQIVTKLQAFQNKTIIVFKQSSGGDPEIVMHSFNCGGEPFYCNSTK 387
Qy      332 LFNSTWNN-----TIGPNNNTGTTILPCRKQIINRMQEVGKAMYAPPIRGQIRCSSNIT 386
Db      388 LFNSTWNNSTWNTGTEGSSNTEGTTILPCRKQIINRMQEVGKAMYAPPIRGQIRCSSNIT 447
Qy      387 GLLLTRDGGKEISNT--TEIFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKARRVVQ 444
Db      448 GLLLTRDGG---SNTSGTEIFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKARRVVQ 504
Qy      445 REKRAVTLGAMFGLGFLGAAGSTWGARSLLTVQARQLLSGIVQOQNNLLRAIEAQHLLQ 504
Db      505 REKRAVGIGAVFLGFLGAAGSTWGAAASMTLTVQARQLLSGIVQOQNNLLRAIEAQHLLQ 564
Qy      505 LTWVGIK-LOARVLAVERYLKQDOLLGWCSCGLICTTAVPWNASNSKSLDOINNNMT 563
Db      565 LTWVGIKQLQARVLAVERYLKQDOLLGWCSCGLICTTAVPWNASNSKSLDKIWNMT 624
Qy      564 WMEWERIDNNTLIYTLIEESQKNEQELLEDKWSLNNWFDISKWLWYIK 619
Db      625 WMEWERINNTSYLIYTLIEESQKNEQELLEDKWSLNNWFDITKWLWYIK 680

RESULT 7
Q70150
ID Q70150 PRELIMINARY; PRT; 843 AA.
AC Q70150;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=95194694; PubMed=7888189;
RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,
RA McCutchan F.B., Bradaç J.A., Sharp P.M., Hahn B.H.;
RA "Genetic variation of HIV type 1 in four World Health Organization-
RT sponsored vaccine evaluation sites: generation of functional envelope
RT (glycoprotein 160) clones representative of sequence subtypes A, B, C,
RT and E. WHO Network for HIV Isolation and Characterization.";
RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=95194691; PubMed=7545977;
RA WHO Global Programme on AIDS;
RA "HIV type 1 variation in World Health Organization-sponsored vaccine
RT evaluation sites: genetic screening, sequence analysis, and
RT preliminary biological characterization of selected viral strains. WHO
RT Network for HIV Isolation and Characterization.";
```

```
RL AIDS Res. Hum. Retroviruses 10:1327-1343 (1994).
RP [3]
RN SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
RA Hahn B.H.;
RT "Molecular cloning and analysis of functional envelope genes from
RT human immunodeficiency virus type 1 sequence subtypes A through G. The
RT WHO and NIAID Networks for HIV Isolation and Characterization.";
RL J. Virol. 70:1651-1657 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RC STRAIN=E.E.;
RA Allen E.E.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U08801; AB05185.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON TER 1
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 843 AA; 95679 MW; 14DB49A6D2F5FC4 CRC64;

Query Match 87.9%; Score 2925; DB 15; Length 843;
Best Local Similarity 84.9%; Pred. No. 6.5e-230;
Matches 551; Conservative 25; Mismatches 39; Indels 34; Gaps 5;

Qy 1 SAVEKLVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN 60
Db 26 SAAEKLVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPOEIVLGN 85
Qy 61 VTENFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATKSSNWKEM 120
Db 86 VTENFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATKSSNWKEM 141
Qy 121 DRGEIKNCSPKVGAG-----KLINCNTSVITQACPKVSF 154
Db 142 EGGEIKNCSEFNITTSIKTKVDYALFYKLDVVPIDNDNTSYRLINCNTSVITQACPKVSF 201
Qy 155 EPIPIHYCAPAGFAILLKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNGLAEEGV 214
Db 202 EPIPIHYCTPAGFAILLKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNGLAEEV 261
Qy 215 VIRSENFDTNAKTIIVOLKESVEINCTRPNNTRKSIITIGPRAFYATGDIIGDIRQAH 274
Db 262 VIRSSNFTDNARVIIVOLKESVEINCTRPNNTRKSIHILGPRAWYTTGQIIGDIRQAH 321
Qy 275 NISGEKNWNTLKOIVTKLOAQFGNKTIVFKQSGGDPPEIVMHSFNCGGEFFYCNSTOLF 334
Db 322 NLSSTKWNNTLROITKLEQFGNKTIVFNQSGGDPPEIVMHSFNCGGEFFYCNSTOLF 381
Qy 335 STWNNTIGPNT--NGTITLPCRIKOIINRQWGVKAMYPPIRGQIRCSNITGLLLTR 392
Db 382 STWNNTIGPNT--NGTITLPCRIKOIINRQWGVKAMYPPIRGQIRCSNITGLLLTR 441
Qy 393 DGGKEISNTTEIPRPGGDMRDNRSELYKVKVIEPLGVAPTKAKRRVVQREKRAV-T 451
Db 442 DGGENSEKTTETPRPGGDMRDNRSELYKVKVIEPLGVAPTKAKRRVVQREKRAVGT 501
Qy 452 LGAMFLGFLGAAGSTMGARSITLTVOARQLLSGIVQOQNLLRAIEAQHLLQITVWGIK 511
Db 502 IGAMFLGFLGTAGSTMGAASITLTVOARQLLSGIVQOQNLLRAIEAQHLLQITVWGIK 561
Qy 512 -LQARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASWNSKSLDQIWNNTWMEWERE 570
Db 562 -LQARVLAVERYLKDQQLLGWCGSGKLICTTAVPWNASWNSKSLDQIWNNTWMEWERE 621
Qy 571 IDNYTNLIYTLIBESQKQKNEQLELDKQASLWNNWFDISKWLWYIK 619
Db 622 IDNYTREIYTLIBESQKQKNEQLELDKQASLWNNWFDITKWLWYIK 670

RESULT 8
O40222 PRELIMINARY; PRT; 854 AA.
AC O40222;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RX MEDLINE=96432129; PubMed=8835195;
RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA Martin M.A., Peden K.W.;
RT "Construction and characterization of a stable full-length macrophage-
RT tropic HIV type 1 molecular clone that directs the production of high
RT titers of progeny virions.";
RL AIDS Res. Hum. Retroviruses 12:191-194 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA Martin M.A., Peden K.W.C.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004394; AB64170.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97291 MW; 06C45E69103C6C12 CRC64;

Query Match 87.9%; Score 2924.5; DB 15; Length 854;
Best Local Similarity 85.1%; Pred. No. 7.2e-230;
Matches 558; Conservative 17; Mismatches 42; Indels 39; Gaps 7;

Qy 1 SAVEKLVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPOEIVLEN 60
Db 28 SAVENLVTVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPOEVLEN 87
Qy 61 VTENFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATKSSNWKEM 120
Db 88 VTENFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNAATKSSNWKEM 145
Qy 121 DRGEIKNCSPKVGAG-----KLINCNTSVITQACPKVS 153
Db 146 MRGEIKNCSEFNITTSIRDKYKDYALFYKLDVVPIDNDNTSYRLINCNTSVITQACPKVS 205
Qy 154 FEPIPIHYCAPAGFAILLKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNGLAEEG 213
Db 206 FEPIPIHYCTPAGFAILLKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNGLAEE 265
Qy 214 VIRSENFDTNAKTIIVOLKESVEINCTRPNNTRKSIITIGPRAFYATGDIIGDIRQAH 273
Db 266 VIRSSNFTDNARVIIVOLKESVEINCTRPNNTRKSIHILGPRAWYTTGQIIGDIRQAH 325
Qy 274 CNISGEKNWNTLKOIVTKLOAQFG-NKTIIVFKQSGGDPPEIVMHSFNCGGEFFYCNSTOL 332
Db 332 CNISGEKNWNTLKOIVTKLOAQFG-NKTIIVFKQSGGDPPEIVMHSFNCGGEFFYCNSTOL 392
```

```
Db 326 CNISRTKWNNTLQIATKKEQFGNNKTIIVFNQSSGGDPEIVMHSFNCGGEFFYCNSTQL 385
Qy 333 FNSTWN-----NTGPNNTNG--TITLPCRKIKIINRWQEVGKAMYAPPIRGIRCSNI 385
Db 386 FNSTWPNFGTWNLTQSTNGTGENDTITLPCRKIKIINRWQEVGKAMYAPPIRGIRCSNI 445
Qy 386 TGLLLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKRRVVQR 445
Db 446 TGLLLTRDGGNNHNDTETFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKRRVVQR 505
Qy 446 EKRAV-TLGMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQOQNLLRAIEAQHLLQ 504
Db 506 EKRAVGITGAMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQOQNLLRAIEAQHLLQ 565
Qy 505 LTWVGIK-LQARVLAVRYLKDDQLLGIWGCSSGKLICTTAVPWNASWSNKSLOIWNMT 563
Db 566 LTWVGIRQLQARVLAVRYLRDQQLLGIWGCSSGKLICTTAVPWNASWSNKTLDIWNMT 625
Qy 564 WMEREIDVNTLIYTLIEESQOQKNEQELLELDKWSLWNNWFDISKWLWYIK 619
Db 626 WMEREIDVNTGLIYTLIEESQOQKNEQELLELDKWSLWNNWFDITWNLWYIK 681

RESULT 9
Q9PXW7 PRELIMINARY; PRT; 846 AA.
AC Q9PXW7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP160, envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93189881; PubMed=8446773;
RA Schmidtmyerova H., Gayet O., Guettari N., Bolmont C., Hirsch I.,
RA Chermann J.C.;
RT "Characterization of HIV1-PAR, a macrophage-tropic strain: cell
RT tropism, virus/cell entry and nucleotide sequence of the envelope
RT glycoprotein.";
RL Res. Virol. 144:21-26(1993).
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Esv GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 846 AA; 96490 MW; 3D561D5735F90C84 CRC64;

Query Match 87.6%; Score 2914.5; DB 15; Length 846;
Best Local Similarity 84.7%; Pred. No. 4.7e-229;
Matches 548; Conservative 27; Mismatches 43; Indels 29; Gaps 3;

Qy 1 SAVEKLWTVYGVGVVWKEATTLTFCASDAQYDTEVHNWATHACVPTDPNPQIVLEN 60
Db 28 SATDKLWTVYGVGVVWKEATTLTFCASDAQYDTEVHNWATHACVPTDPNPQIVLEN 87
Qy 61 VTENFMKNNMVEQMEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNATNTKSNWKEM 120
Db 88 VTEYFNMKNNMVEQMEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNATNTSRNGGM 147
Qy 121 DRGEIKNSFKVGAG-----KLINCNTSVITQACPKVS 153
Db 148 EKGEIKNSFNITTSIRDVKQYALFYKLDVVPIDNNSTRYRLISCNSTVITQACPKVT 207
Qy 154 FEPIPIHCAPAGFAILCKNDKKNFGSGPCTNVTQCTHGIRPVVSTQLLNGSLAEEG 213
```

```
Db 208 FEPIPIHCAPAGFAILCKKDKKFDGKGPCTNVTQCTHGIRPVVSTQLLNGSLAEE 267
Qy 214 VVIRSEFTDNAKTIIVQLKESVEICTRPNNNTRKSITIGPGRAFATGDIIGDIRQAH 273
Db 268 VVIRSEFTDNAKTIIVQLKESVEICTRPNNNTRKSITIGPGRAFATGDIIGDIRQAH 327
Qy 274 CNISGKWNNTLQIATKKEQFGNNKTIIVFNQSSGGDPEIVMHSFNCGGEFFYCNSTQLF 333
Db 328 CTISKTWENTFQIIVKRLREQYKNTIIVFNQSSGGDPEIVTHSFNCGGFFYCNSTQLF 387
Qy 334 NSTWNTTGNNTNGTITLPCRKIKIINRWQEVGKAMYAPPIRGIRCSNITGLLITRD 393
Db 388 NSTWNTTGNNTTERTITLPCRKIKIINRWQEVGKAMYAPPIRGIRCSNITGLLITRD 447
Qy 394 GKKEISNTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKRRVVQREKRAVTLG 453
Db 448 GGNNNNGTIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKRRVVQREKRAVTLG 507
Qy 454 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWVGIK-L 512
Db 508 AMFLGFLGAAGSTMGARSLLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWVGIKQL 567
Qy 513 QARVLAVRYLKDDQLLGIWGCSSGKLICTTAVPWNASWSNKSLOIWNMTWMEWEREID 572
Db 568 QARVLAVRYLKDDQLLGIWGCSSGKLICTTAVPWNASWSNKSLOIWNMTWMEWEREID 626
Qy 573 NYTNLIYTLIEESQOQKNEQELLELDKWSLWNNWFDISKWLWYIK 619
Db 627 NYTNLIYTLIEESQOQKNEQELLELDKWSLWNNWFDITRNLWYIK 673

RESULT 10
Q75760 PRELIMINARY; PRT; 847 AA.
AC Q75760;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=JRFL;
RX MEDLINE=87206194; PubMed=3646751;
RA Koyanagi Y., Miles S., Mitsuyasu R.T., Merrill J.E., Vinters H.V.,
RA Chen I.S.;
RT "Dual infection of the central nervous system by AIDS viruses with
RT distinct cellular tropisms.";
RL Science 236:819-822(1987).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=JRFL;
RX MEDLINE=91043044; PubMed=2172833;
RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K.,
RA Zack J.A., Chen I.S.;
RT "HIV-1 tropism for mononuclear phagocytes can be determined by regions
RT of gp120 outside the CD4-binding domain.";
RL Nature 348:69-73(1990).
[3]
RP SEQUENCE FROM N.A.
RX STRAIN=JRFL;
RX MEDLINE=92092169; PubMed=1684385;
RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.;
RT "HIV-1 env sequence variation in brain tissue of patients with AIDS-
RT related neurologic disease.";
RL J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).
[4]
RP SEQUENCE FROM N.A.
RX STRAIN=JRFL;
RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.,
```



RA Koyanagi Y., Namazie A., Zhao J., Diagne A., Idler K.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63632; AAB05604.1; -.  
DR PIR; S13289; S13289.  
DR PIR; T09448; T09448.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR000328; Env Gp41.  
DR InterPro; IPR000777; Gp120.  
DR Pfam; PF00516; Gp120; 1.  
DR Pfam; PF00517; Gp41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 847 AA; 96160 MW; 022D5F24E04FB29F CRC64;

Query Match 87.3%; Score 2906; DB 15; Length 847;  
Best Local Similarity 85.4%; Pred. No. 2.3e-228;  
Matches 554; Conservative 25; Mismatches 38; Indels 32; Gaps 6;  
QY 1 SAVEKLWTVVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEIIVLEN 60  
DB 28 SAVEKLWTVVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 87  
QY 61 VTENFNNKNNWVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTNKNATNTKSSNWKEM 120  
DB 88 VTEHFNWNNWVQMQEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTW 146  
QY 121 DRGKIKNSPKVAG-----KLIINCNTSVITQACPVS 153  
DB 147 ERGEIKNSFNITTSIRDEVOKEVALFYKLDVDPIDNNNTSYRLISCDTSVITQACPIS 206  
QY 154 FEPIPIHYCAPAGFAILKCNCKKFNKSGPCTNVTQCTHGIRPVVSTQLLNGSLAE 213  
DB 207 FEPIPIHYCAPAGFAILKCNCKTENGKPCNVSTVQCTHGIRPVVSTQLLNGSLAE 266  
QY 214 VVIRSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITGPGRAFATGDIIGDIA 273  
DB 267 WIRSDNFTNAKTIIVQLKESVEINCTRPNNTRKSIHIGPGRAFYTGTGIIIGDIA 326  
QY 274 CNTSGEKWNTLKOIVTKLQAOFGNKTIVFKQSGGDPPEIVMHSFNCGGEFFVCNSTOLF 333  
DB 327 CNTSRAKWNLTQKIVTKLQREQFNKTIIVFNHSGGDPPEIVMHSFNCGGEFFVCNSTOLF 386  
QY 334 NSTW--NNTIGPNNTG--TITLPCRKQIINRWQEVGKAMYAPPPIRGQIRCSSNITGLLT 391  
DB 387 NSTWNNTEGNSNTEGNTITLPCRKQIINRWQEVGKAMYAPPPIRGQIRCSSNITGLLT 446  
QY 392 RDGKEISNTEIPRPGGDMRDNRSELKYKVKVKEIPLGVAPTAKRRVQREKAVT 451  
DB 447 RDGGIN--ENGTEIPRPGGDMRDNRSELKYKVKVKEIPLGVAPTAKRRVQREKAVG 505  
QY 452 LGAMFLGFLGAGSTMGARSITLTQARQLLSGIVQOONLLRAIEAQOHLQLTWYG 511  
DB 506 IGAVFLGFLGAGSTMGAASTLTQARQLLSGIVQOONLLRAIEAQOHLQLTWYG 565  
QY 512 -LQARVLAVERYLKQDQLLGIWCGSGKLICTTAVPNNASNSKSLDOIWNMTWMEW 570  
DB 566 QLQARVLAVERYLDQQLLGIWCGSGKLICTTAVPNNASNSKSLDRIWNMTWMEW 625  
QY 571 IDNVTNLIYTLIBESNQEKNEQELLELDKWSLWNNFDISKWLWYIK 619  
DB 626 IDNVTSEIYTLIBESNQEKNEQELLELDKWSLWNNFDTIKWLWYIK 674

RESULT 11  
O41883  
ID O41883 PRELIMINARY; PRT; 852 AA.  
AC O41883;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.

GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SC14.3;  
RX MEDLINE=98178716; PubMed=9519894;  
RA McCutchan F.E.; Sanders-Buell E.; Salminen M.O.; Carr J.K.;  
RA Sheppard W.H.;  
RT "diversity of the human immunodeficiency virus type 1 envelope  
glycoprotein in San Francisco Men's Health Study participants.";  
RL AIDS Res. Hum. Retroviruses 14:329-337(1998).  
DR EMBL; U90934; AAC59271.1; -.  
DR PIR; A53591; A53591.  
DR PIR; T12016; T12016.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR000328; Env Gp41.  
DR InterPro; IPR000777; Gp120.  
DR Pfam; PF00516; Gp120; 1.  
DR Pfam; PF00517; Gp41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 852 AA; 96549 MW; 609952E37D7CF064 CRC64;  
Query Match 87.2%; Score 2900.5; DB 15; Length 852;  
Best Local Similarity 84.3%; Pred. No. 6.6e-228;  
Matches 552; Conservative 32; Mismatches 32; Indels 39; Gaps 7;  
QY 1 SAVEKLWTVVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEIIVLEN 60  
DB 28 SAAEQLRVTVVYGVVPVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 87  
QY 61 VTENFNNKNNWVEQMHEDIISLDQSLKPCVKLTPLCVTLHCTN-LKNAATNTKSSNKE 119  
DB 88 VTENFNNKNNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDYLNRDNTTSSNNGG 147  
QY 120 MDRGEIKNSCFK-----GAGKLIINCNTSVITQACP 152  
DB 148 MEGGEIKNSCFNITTRIGNVKQKEYALFYKLDVVPIDNTTTSYRLINCNTSVITQACP 207  
QY 153 SFPEPIPIHYCAPAGFAILKCNCKKFNKSGPCTNVTQCTHGIRPVVSTQLLNGSLAE 212  
DB 208 SFPEPIPIHYCTPAGFALLKCNCKKFNKSGPCTNVTQCTHGIRPVVSTQLLNGSLAE 267  
QY 213 GVVISSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIITGPGRAFATGDIIGDIA 272  
DB 268 EVVISSENFDTNAKTIIVQLKESVEINCTRPNNTRKSIHIGPGQALYATGAIIGDIA 327  
QY 273 HCNIISGEKWNLTQKIVTKLQAOFGNKTIVFKQSGGDPPEIVMHSFNCGGEFFVCNSTOLF 332  
DB 328 HCNISRAKWNLTQKIVTKLQVQFGNKTIIIFNOSGGDPPEIVMHSFNCGGEFFVCNTTKL 387  
QY 333 FNSTW--NNTIGPNNTG-----NCTITLPCRKQIINRWQEVGKAMYAPPPIRGQIRCSSNI 385  
DB 388 FNSTWNNFNTW--NDTETEGNCTITLPCRKQIINRWQEVGKAMYAPPPIRGQIRCSSNI 445  
QY 386 TGLLLTRDGGKEISNTEIPRPGGDMRDNRSELKYKVKVKEIPLGVAPTAKRRVQ 445  
DB 446 TGLLLTRDGGTN--NSTNETPRPGGDMRDNRSELKYKVKVKEIPLGVAPTAKRRVQ 504  
QY 446 EKEAVTLGAMFLGFLGAGSTMGARSITLTQARQLLSGIVQOONLLRAIEAQOHLQL 505  
DB 505 EKEAVGIGALFLGFLGAGSTMGAAVTLTQARQLLSGIVQOONLLRAIEAQOHLQL 564  
QY 506 TWGKIK-LQARVLAVERYLKQDQLLGIWCGSGKLICTTAVPNNASNSKSLDOIWNMTW 564  
DB 565 TWGKIKQLQARVLAVERYLDQQLLGIWCGSGKLICTTAVPNNASNSKSLDKIWNMTW 624  
QY 565 MEWERIDNVTNLIYTLIBESNQEKNEQELLELDKWSLWNNFDTIKWLWYIK 619



Db 328 NLSEAKNHTLEQIAKLRQFGNKTIVFNQSSGDDPEIVMHSFNCGEFFYCNSTKLEN 387  
Qy 334 -----NSTWNNTIGPNNTNGTITLPCRIKQIINRWQEVGKAMYAPPPIRGQIRCSNITGL 388  
Db 388 RTWSVNSTWMDTEGVNNTGENITLPCRIKQIINRWQEVGKAMYAPPPIKGQIRCSNITGL 447  
Qy 389 LLTRDGGKEISNTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARRVVOREKR 448  
Db 448 LLTRDGGNS--NNEETETFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARRVVOREKR 506  
Qy 449 AVTLGAMFLGFLGAAGSTMARSLLTVQARQLLSGIVQOQNNLLRAIEAQHLLQTLTV 508  
Db 507 AVGLGAVFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQOQNNLLRAIEAQHLLQTLTV 566  
Qy 509 GIK--LOARVLAVERYLKDQQLLGGWCGSGKLICTTAVPWNASWSNKSLEKIDWNNMTWMEW 567  
Db 567 GIKQLQARVLAVERYLKDQQLLGGWCGSGKLICTTAVPWNASWSNKSLEKIDWNNMTWMEW 626  
Qy 568 EREIDNTNLIYTLIEESNQOQEKNEQELLELDKWSLNNWFDISKWLWYIK 619  
Db 627 DREINNTYSLIYTLIEESNQOQEKNEQELLELDKWSLNNWFDITNLWYIK 678  
  
RESULT 14  
ID O92762 PRELIMINARY; PRT; 854 AA.  
AC O92762;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RC STRAIN=SPMHS3.3;  
RX MEDLINE=98178716; PubMed=9519894;  
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,  
RA Sheppard W.H.;  
RT "Diversity of the human immunodeficiency virus type 1 envelope  
glycoprotein in San Francisco Men's Health Study participants.";  
RL AIDS Res. Hum. Retroviruses 14:329-337(1998).  
DR EMBL; AF025754; AAC40591.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env GP41.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
SQ SEQUENCE 854 AA; 97287 MW; AFD8A82B62CSB27D CRC64;  
  
Query Match 87.0%; Score 2894; DB 15; Length 854;  
Best Local Similarity 83.6%; Pred. No. 2.2e-227; Indels 34; Gaps 6;  
Matches 546; Conservative 29; Mismatches 44;  
  
Qy 1 SAVEKLWTVYGVVWKEATTTILFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60  
Db 29 SATEKLWTVYGVVWKEATTTILFCASDAKAYDTEAHNIWATHACVPTDPNPQEVLEN 88  
  
Qy 61 VTENFNKNNWVQMHEDIISLWQSLKPCVKLTPLCVTLHCTN-LKANTNTKSNWKE 119  
Db 89 VTENFNKNNWVQMHEDIISLWQSLKPCVKLTPLCVTLKCTNDLKNNTTNTNNSWEK 148  
  
Qy 120 MDRGEIKNCSFKVCGAG-----KLINCNTSVLTQACPKV 152  
Db 149 METGEIKNCSFNITNRRGKMQKEYALFYKLDVVVSINDNTSYRLISNTSVITQACPKV 208

Qy 153 SPFPIPIHYCAPAGFAILLKNDKKFNKSGPCTNVTSTVQCTHGIRPVVSTQLLINGSLAEE 212  
Db 209 SFOPIPIHYCAPAGFAILLKNDKKFKSGKPCINVTSTVQCTHGIRPVVSTQLLINGSLAEE 268  
Qy 213 GVVRSNFTDNAKTIIVQLKESVEINCTRPNNNTRKSIITIGPGRAFVATGDIIGDIOQA 272  
Db 269 EVVRSNFTDNAKTIIVQLKEPVEINCTRPNNNTRKSIIPIGPGRAFYTTTGIIIGNIOQA 328  
Qy 273 HCNISGKWNNTLQIIVTKLQAFGNKTIIVFKOSSGDDPEIVMHSFNCGEFFYCNSTQL 332  
Db 329 HCNISRAKWNNTLQIIVIELKEHFKNKTIVFNHSSGDDPEIVMHSFNCGEFFYCNSTQL 388  
Qy 333 FNSTWNT--IGPNNTNG--TITLPCRIKQIINRWQEVGKAMYAPPPIRGQIRCSNITGL 388  
Db 389 FSSWTNNTSEGVNNTTEGNDTIIILLCRIKQIIVNMWQEVGKAMYAPPPIRGQIRCSNITGL 448  
Qy 389 LLTRDGGKEI-SNTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARRVVOREK 447  
Db 449 LLTRDGGNSGSKAEVFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARRVVOREK 508  
Qy 448 RAVTLGAMFLGFLGAAGSTMARSLLTVQARQLLSGIVQOQNNLLRAIEAQHLLQTLTV 507  
Db 509 RAAGLVNMFGLGFLGAAGSTMGAASMTLTVQARQLLSGIVQOQNNLLRAIEAQHLLQTLTV 568  
Qy 508 WGIK-LOARVLAVERYLKDQQLLGGWCGSGKLICTTAVPWNASWSNKSLEKIDWNNMTWME 566  
Db 569 WGIKQLQARVLAVERYLKDQQLLGGWCGSGKLICTTAVPWNNTSNKSLDRINWNNMTWME 628  
Qy 567 WESEIDNTNLIYTLIEESNQOQEKNEQELLELDKWSLNNWFDISKWLWYIK 619  
Db 629 WEKEIDNTYSLIYTLIEESNQOQEKNEQELLELDKWSLNNWFDITNLWYIK 681  
  
RESULT 15  
ID O56108 PRELIMINARY; PRT; 853 AA.  
AC O56108;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX STRAIN=SPMHS3.1;  
RX MEDLINE=98178716; PubMed=9519894;  
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,  
RA Sheppard W.H.;  
RT "Diversity of the human immunodeficiency virus type 1 envelope  
glycoprotein in San Francisco Men's Health Study participants.";  
RL AIDS Res. Hum. Retroviruses 14:329-337(1998).  
DR EMBL; AF025751; AAC40589.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 853 AA; 96830 MW; 5C8BA4418F8AEC47 CRC64;  
  
Query Match 87.0%; Score 2893; DB 15; Length 853;  
Best Local Similarity 84.2%; Pred. No. 2.7e-227;  
Matches 550; Conservative 24; Mismatches 45; Indels 34; Gaps 4;  
  
Qy 1 SAVEKLWTVYGVVWKEATTTILFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 60  
Db 28 SAAEKLWTVYGVVWKEATTTILFCASDAKAYDTEVHNWATHACVPTDPNPQEVLEN 87

Qy	61	VTENFMKNNMVEQMHEDIISLWDSIKPCVKLTPLCVTLHCT-NLKNATNTKSSNWKE	119
Db	88	VTENFMKNNMVEQMHEDIISLWDSIKPCVKLTPLCVTLNRTDDDKNATNTNSSGEM	147
Qy	120	MDRGEIKNCSFKVGAG-----KLINCNTSVITQACP	150
Db	148	EMRGEIKNCSFNITTSIRDKYKEYALFYKLDVVPINNNTTTSYRLINCNTSVITQACP	207
Qy	151	KVSFEPIPIHYCAPAGFAILKCNCKFNGSPCTNVSTVQCTHGIRPVVSTQLLNGSLA	210
Db	208	KVSFEPIPIHYCTPAGFAILKCKKFKNGTGPCTSVSTVQCTHGIRPVVSTQLLNGSLA	267
Qy	211	EEGVVIRSENETDNAKTIIVOLKESVEINCTRPNNTRKSIITIGPRAFYATGDIIGDIR	270
Db	268	EEVVVIRSENETDNAKIIIVQLNESVESNCTRPNNTRKSIPIGPRAFYTTGIIIGDIR	327
Qy	271	QAHNCNISGEKNNTLKQIVTKLOAQFGNKTIVFKQSSGGDPEIVMHSFNCGGEFFYCNS	330
Db	328	QAHNCNLSRTKWDNALQIACKLGHQFGTKTIVFNQSSGGDPEIVMHTFNCGGEFFYCNS	387
Qy	331	QLFNSTWNTTGPNNNT---NGTITLPCRKQIINRWQEVGKAMYAPPPIRGQIRCSSNITG	387
Db	388	QLFNSTWNTTRESNNTGNGTITLPCRKQIINRWQGVGKAMYAPPPIRGQIRCSSNITG	447
Qy	388	LLLTRDGGKEISNTTEIFRPGGDMRDNWRSELYKYVVKIEPLGVAPTAKRRVVOREK	447
Db	448	LLLTRDGGNSESNETETFRPGGDMRDNWRSELYKYVVKIEPLGVAPTAKRRVVOREK	507
Qy	448	RAVTLGAMFLGFLGAAGSTMGARSILTVQARQLLSGIVQQONNLLRAIEAQHLLQITV	507
Db	508	RAVGIGAVFLGFLGAGSTMGAASWTLTVQARLLLSGIVQQONNLLRAIEAQHLLQITV	567
Qy	508	WGIK-LQARVLAVERYLKDQQLLGWCSGKLICTTAVPMNASWSNKSOLDQIWNNTWME	566
Db	568	WGIKQARVLAVERYLKDQQLLGWCSGKLICTTTVPMNASWSNKSOLDKIWNNTWME	627
Qy	567	WEREIDNYTNIYTLIEESQOQEKNELELDKASLWNNWFDISKWLWYIK	619
Db	628	WEREIDNYTSLIYTLIEESQOQEKNELELDKASLWNNWFDITKWLWYIK	680

Search completed: February 25, 2004, 14:18:37  
 Job time : 41.1462 secs